File No. 3057.1B

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:	)
	)
PETER FIEKOWSKY et al	) Examiner: Edward Raymond
	)
Serial No. 10/648,819	) Art Unit: 1645
	)
	) Response to Office Action
	) Dated August 25, 2004, And
Filed: August 25, 2003	) Request for Declaration of Interference
	) with a Patent under 37 CFR §1.607 -
For: SCANNED IMAGE ALIGNMENT SYSTEMS AND METHODS	) Expedited Handling
	)

COMMISSIONER OF PATENTS
Alexandria, VA 22313

Sir:

REMARKS

This paper is in response to the Office Action dated August 25, 2004 setting a one month time limit to provide the support for the claims copied from two U.S. patents. This paper provides that support and the other formalities required to declare an interference between the present application and the two patents.

In accordance with 37 C.F.R. §1.607, Applicants hereby request an interference with United States Patent Nos. 6,591,196 and 6,768,820 which issued to Yakhini et al. on July 8, 2003 and July 27, 2004 respectively. To facilitate consideration of this request, Applicants attach a proposed PTO-850 "Interference Initial Memorandum" outlining the requested interference.

Applicants herein comply with the provisions of 37 C.F.R. §1.607, which requires the following:

(1) Identify the patent;

(2) Present a proposed count;

(3) Identify at least one claim in the patent corresponding to the proposed count;

(4) Present at least one claim corresponding to the proposed count or identify at least one claim already pending in its application that corresponds to the proposed count, and, if any claim of the patent or application identified as corresponding to the proposed count does not correspond exactly to the proposed count, explain why each such claim corresponds to the proposed count; and

(5) Apply the terms of any application claim,

(i) Identified as corresponding to the count, and

(ii) Not previously in the application to the disclosure of the application.

(6) Explain how the requirements of 35 U.S.C. § 135(b) are met, if the application claim identified as corresponding to the proposed count was not present in the application until more than one year after the issue date of the patent.

37 C.F.R. §1.607(a) (1) - Identification of Involved Patents

In accordance with 37 C.F.R. §1.607(a) (1), Applicants identify U.S. Patent Nos. 6,591,196 and to Yakhini et al. ("the '196 patent") and 6,768,820 ("the '820 patent") which issued to Yakhini et al. on July 8, 2003 and July 27, 2004 respectively.

37 C.F.R. §1.607(a) (2) - Proposed Count

In accordance with 37 C.F.R. §1.607(a) (2), Applicants propose a count defined as follows:

Claim 14 of the '196 patent

or

Applicants' Claim 46

or

Claim 1 of the '820 patent

or

Applicants claim 64

Claim 14 of the '196 patent reads as follows:

A system for automated extraction of data from a molecular array having features arranged in a regular pattern, the system comprising:

a scanning component that produces images of the molecular array representing intensities of data signals emitted from discrete positions on a surface of the molecular array;

a computer program that processes the images of the molecular array produced by the scanning component to index features in the images of the molecular array corresponding to molecules bound to features of the molecular array and that extracts data from the indexed features within images of the molecular array; and

a computer for executing the computer program.

Applicants' claim 46 reads as follow:

A system for automated extraction of data from a molecular array having features arranged in a regular pattern, the system comprising:

a scanning component that produces images of the molecular array representing intensities of data signals emitted from discrete positions on a surface of the molecular array;

a computer program that processes the images of the molecular array produced by the scanning component to index features in the images of the molecular array corresponding to molecules bound to features of the molecular array and that extracts data from the indexed features within images of the molecular array; and

a computer for executing the computer program.

Claim 1 of '820 reads as follows:

1. A method for evaluating an orientation of a molecular array having features arranged in a pattern, the method comprising:

(a) receiving an image of the molecular array produced by scanning the molecular array to determine data signals emanating from discrete positions on a surface of the molecular array;

(b) calculating an actual result of a function on pixels of the image lying in a second pattern;

(c) comparing the result of step (b) with an expected result which would be obtained if the second pattern had a predetermined orientation on the array; and

(d) when the results of the comparison in step (c) are outside a predetermined difference, then altering the orientation of the second pattern on the array and repeating steps (b) and (c), and repeating the foregoing as needed until the results of the comparison are within the predetermined difference.

Applicants claim 64 reads as follows:

1. A method for evaluating an orientation of a molecular array having features arranged in a pattern, the method comprising:

(a) receiving an image of the molecular array produced by scanning the molecular array to determine data signals emanating from discrete positions on a surface of the molecular array;

(b) calculating an actual result of a function on pixels of the image lying in a second pattern;

(c) comparing the result of step (b) with an expected result which would be obtained if the second pattern had a predetermined orientation on the array; and

(d) when the results of the comparison in step (c) are outside a predetermined difference, then altering the orientation of the second pattern on the array and repeating steps (b) and (c), and

repeating the foregoing as needed until the results of the comparison are within the predetermined difference.

37 C.F.R. §1.607(a) (3) - Patent Claims Corresponding to the Proposed Count

In accordance with 37 C.F.R. §1.607(a) (3), Applicants identify claims 1-18 of the '196 patent as corresponding to the proposed count. Claim 14 is expressly recited in the definition of the proposed count and claims 1-13 and 15-18 define the same patentable invention as the proposed count, as explained below. Applicants also identify claims 1-5 of the '820 patent as corresponding to the proposed count. Claim 1 is expressly recited in the definition of the proposed count and claims 2-5 define the same patentable invention as the proposed count, as explained below.

37 C.F.R. §1.607(a) (4) - Application Claims Corresponding to the Proposed Count

In accordance with 37 C.F.R. §1.607(a) (4), Applicants identify Applicants' claims 46-61 and 64-65 as corresponding to the proposed count. Applicants' claims 59 and 64 are expressly recited in the definition of the proposed count and Applicants' claims 46-58, 60-61 and 65 define the same patentable invention as the proposed count, as explained below.

37 C.F.R. §1.607(a) (5) – Applying Terms of Application Claims to the Disclosure

Applicants' claims 46-61 were copied from '196 and are hereby presented in the application which have not yet been substantively examined. Accordingly, Applicants apply the terms of claims 46-61 (corresponding to the '196 claim) to the disclosure of the application in the table set forth below.

Claim 46 Independent	*The Present Application*
A method, embodied in a computer program, for automated extraction data from a molecular array having features arranged in a regular pattern, the method comprising:	1) Exemplary computer code for processing scanned images executed by analysis system 126 is presented in the "Software Appendix" pages 23-25. to "computer program for automated extraction data" Fig 3, and page 9, line 25 through page 10, line 10 describe a computerized system for forming and analyzing arrays of biological materials. Also, the analysis system is further described in page 11, lines 20-25. Further support for computer program for extraction of data from a molecular array. Arrays of polymers or chips are described in page 10, lines 20-22, where the chip includes a first pattern of nucleic acid sequences (page 4, lines 17-19). Analogous to "molecular array having features arranged in a regular pattern"
receiving a number of images of the molecular array, each produced by scanning the molecular array to determine intensities of data signals emanating from discrete positions on a surface of the molecular array;	2) Figure 1, and page 11, lines 10-19 describe scanning system 120 that generates image file 124 that indicates "the fluorescence intensity (photon counts or other related measurements such as voltage) as a function of position on the substrate". Relates to receiving an image of the molecular array, produced by scanning. Page 15, lines 13-18 describe step 267 that calculates the relative hybridization intensities for each cell including a calculation of the mean of the pixel values within the cell. "For example, the hybridization intensity for a cell, and therefore the relative hybridization affinity between the probe of the cell and the sample sequence, may be calculated as the mean of the pixel values within the cell" (page 15, lines 14-25). Relevant to "determining intensities of data signals emanating from discrete positions on a surface of the molecular array".
estimating initial positions of selected marker features within an image of the molecular array;	3) Page 16, lines 4-11 describes a "pattern in the scanned image" that includes a checkerboard pattern. ("marker feature"). Figure 11, and page 17, line 22 through page 18, line 2 show the relationship of a selected pixel (pixel 1) with respect to pixels 2-9. Also, Figure 10, and page 18, line 3 through page 19, line 18

	describe method of convolving features in an image that includes calculating the average of pixel intensities of odd numbered pixels and even numbered pixels as represented in Figure 11. Relevant to estimating initial positions of marker feature.
calculating refined positions of the selected marker features within the image of the molecular array;	4) Page 17, lines 3-7, describes "the convolved image is searched for bright areas", that are recognizable patterns in the convolved image. Page 17, lines 11-13; "recognizable pattern 41 in this embodiment is a grid pattern that was generated by the checkerboard pattern when it was convolved with a filter". Search for Bright Areas relevant to the calculation of refined positions of the selected marker features.
using the refined positions of the selected marker features to compute an initial coordinate system for locating features of the molecular array in the number of images of the molecular array;	5) Page 19, lines 20-21: "The recognizable patterns in the convolved image are utilized to align the scanned image". Analogous to computation of initial coordinate system using the refined position of marker features.
using the initial coordinate system to locate positions of strong features within one or more images of the molecular array;	6) Figure 13, and Page 20, line 12 through Page 21, line 7 describe using the aligned grid to calculate an average value of all pixels represented on the vertical and horizontal grid lines. The average value is indicative of the positions of "strong" features. For example, the average value will be high if the grid lines cross one or more high intensity features. Thus, the positions of the "strong features are located using the coordinate system.
refining the positions of strong features within the one or more images of the molecular array by analyzing data signal intensity values in regions of the one or more images of the molecular array that contain the strong features;	7) Figure 13, and Page 20, line 12 through Page 21, line 7 (as with method step 6 above) the method as described above is repeated for all of the representative pixels in each grid line "Then, at a step 553, the system may determine if there are more positions of the grid to analyze. If there are, the position of the grid may be adjusted at a step 555. Therefore, the grids may be moved left or right by one or more pixels before the intensities are summed along grid lines at step 551" Page 20, lines 22-25. Thus the exact positions of the strong features are determined or "refined"
using the refined positions of strong features in the one or more images of the molecular array to calculate a refined coordinate system to locate positions of weak features within the number of images of the molecular array;	8) Page 20, line 25 through Page 21, line 2: "Once all the positions of the grid have been analyzed, the system selects a grid position where the pixel intensities (*e.g.* the sum calculated at step 551) are at a minimum. Therefore, if the

	pixel intensities for grid lines are lower at another position, the grid is adjusted accordingly". In other words, the refined coordinate system position is determined based upon the refined position of the strong features. The locations of the weak features are also determined by the refined position of the coordinate system.
using the refined positions of strong features in the one or more images of the molecular array to calculate a refined coordinate system to locate positions of local background regions surrounding all strong and weak features within the number of images of the molecular array; and	9) See method step 8 above and, See page 21, lines 2-3. "This refinement will work well if "the cells are typically separated by a darker area or line." The area between cells is analogous to the "background regions".
extracting data from strong features, and their respective local background regions, within the number of images of the molecular array using the refined positions of strong features within the number of images of the molecular array and extracting data from weak features, and their respective local background regions, within the number of images of the molecular array using locations for the weak features calculated from the refined coordinate system.	Page 15, lines 13-18 describe step 267 that calculates the relative hybridization intensities for each cell including a calculation of the mean of the pixel values within the cell. "For example, the hybridization intensity for a cell, and therefore the relative hybridization affinity between the probe of the cell and the sample sequence, may be calculated as the mean of the pixel values within the cell" (page 15, lines 14-18). Analogous to extracting data using the refined positions of strong features.
Claim 47 Dependent	***The Present Application***
The method of claim 46 wherein data signals emanating from discrete positions on the surface of the molecular array include: fluorescent emission from fluorophores incorporated into molecules bound to features of the molecular array; radiation emitted by radioisotopes incorporated into molecules bound to features of the molecular array; and light emission from chemiluminescent moieties incorporated into molecules bound to features of the molecular array.	Page 11, lines 6-7: "The targets are marked with a label such as a fluorescein label". Page 11, lines 12-19: includes "The output of scanner 120 is an image file(s) 124 indicating, in the case of fluorescein labeled target, the fluorescence intensity (photon counts or other related measurements, such as voltage) as a function of position on the substrate". Page 9, lines 13-18: includes "In a representative embodiment the scanned image files include fluorescence data from a biological array, but the files may also represent other data such as radioactive intensity, light scattering, refractive index, conductivity, electroluminescence, or large molecule detection data". For radiographic labels, see also the patents incorporated by reference at col. 1, lines 27-29. "Such devices are discussed in, for example, U.S. Pat. No. 5,143,854 (Pirrung et al.) incorporated herein by reference for all purposes.

Claim 48 Dependent	*The Present Application*
The method of claim 1 wherein each image of the number of images comprises an array of pixels, each pixel having a data signal intensity value.	Page 15, lines 13-18: includes "the analysis system analyzes the scanned image to calculate the relative hybridization intensities for each cell of interest on the chip"; and "the relative hybridization affinity between the probe of the cell and the sample sequence may be calculated as the mean of the pixel values within the cell. The pixel values may correspond to photon counts from the labeled hybridized sample fragments".
Claim 49 Dependent	*The Present Application*
The method of claim 48 wherein the features of the molecular array are arranged in a rectilinear grid, wherein corner features are selected as marker features, and wherein estimating initial positions of selected marker features within an image of the molecular array further includes: calculating row and column vectors by considering the values of pixels in rows and columns of the image; determining a first and last peak in the row and column vectors; and using pixel coordinates of the first and last peaks in the row vector to determine horizontal coordinates of the corner features and using pixel coordinates of the first and last peaks in the column vector to determine vertical coordinates of the corner features.	Page 16, lines 12-17: includes "A scanned image 301 of a hybridized chip includes an active area 303 where the probes were synthesized. At the corner of the active area is a pattern 305 that is a checkerboard pattern. Typically, the pattern appears at each corner of the active area". See Figs. 7A and 7B. Page 17, line 17 through page 19, line 7; describes convolution method used to identify pattern. Method describes sequentially selecting a pixel and employing the intensity values of neighbor pixels in the calculations where the neighbor pixels may be in various rows and columns as illustrated in Figure 11 (pixels 1-9). Page 19, lines 8-18: describes the determining steps as illustrated in Figures 12A-12D, where a peak of intensity (i.e. when the odds pixels are much brighter than the even pixels, page 18, lines 25-26) may be determined at each iteration of the algorithm described in the calculating steps if the pixel in positioned with the checkerboard feature. A brighter square is generated in the convolved image when the pixel is positioned in the checkerboard pattern and results in the 2X2 square pattern of Figure 9B, thus a first and last peak are determined in row and column vectors. Convolution method averages values of odd numbered pixels and even numbered pixels and assigns an intensity value to the selected pixel based upon the difference between the even and odd averages, the larger the difference the brighter the assigned intensity value. This produces a recognizable pattern (such as

	illustrated in Figure 9B, page 19, line 8, and page 19 line 18). Then "The position of the recognizable pattern in the convolved image may be utilized to align the scanned image, such as by placing a grid over the image"
Claim 59 Independent	***The Present Application***
A system for automated extraction of data from a molecular array having features arranged in a regular pattern, the system comprising: a scanning component that produces images of the molecular array representing intensities of data signals emitted from discrete positions on a surface of the molecular array; a computer program that processes the images of the molecular array produced by the scanning component to index features in the images of the molecular array corresponding to molecules bound to features of the molecular array and that extracts data from the indexed features within images of the molecular array; and a computer for executing the computer program.	Figure 3, and page 9, line 25 through page 10, line 10 describe a computerized system for forming and analyzing arrays of biological materials. Also, the analysis system is further described in page 11, lines 20-25. Arrays of polymers or chips are described in page 10, lines 20-22, where the chip includes a first pattern of nucleic acid sequences (page 4, lines 17-19). Figure 1, and page 11, lines 10-19 describe scanning system 120 that generates image file 124 that indicates "the fluorescence intensity (photon counts or other related measurements such as voltage) as a function of position on the substrate". The scanning system and scanned image are further described in page 15, lines 1-12. Exemplary computer code (i.e. a program) for processing scanned images executed by analysis system 126 is presented in the "Software Appendix", pages 23-25. Page 15, lines 9-12 describe image alignment step 263 as "the scanned image is aligned so that the pixels that correspond to each cell can be identified". Further, page 15, lines 13-18 describe step 267 that calculates the relative hybridization intensities for each cell including a calculation of the mean of the pixel values within the cell. "For example, the hybridization intensity for a cell, and therefore the relative hybridization affinity between the probe of the cell and the sample sequence, may be calculated as the mean of the pixel values within the cell" (page 15, lines 14-25). Figure 3, and page 9, line 25 through pages 10,

	line 10 describe a computerized system for forming and analyzing arrays of biological materials. Also, the analysis system is further described in page 11, lines 20-25.
Claim 60 Dependent	***The Present Application***
The system of claim 14 wherein data signal intensities emanating from discrete positions on the surface of the molecular array include: radiation emitted by radioisotopes incorporated into molecules bound to features of the molecular array; fluorescent emission from fluorophores incorporated into molecules bound to features of the molecular array; and light emission from chemiluminescent moieties incorporated into molecules bound to features of the molecular array.	Page 11, lines 6-7: "The targets are marked with a label such as a fluorescein label" Page 11, lines 12-19: includes "The output of scanner 120 is an image file(s) 124 indicating, in the case of fluorescein labeled target, the fluorescence intensity (photon counts or other related measurements, such as voltage) as a function of position on the substrate" Page 9, lines 13-18: includes "In a representative embodiment the scanned image files include fluorescence data from a biological array, but the files may also represent other data such as radioactive intensity, light scattering, refractive index, conductivity, electroluminescence, or large molecule detection data"
Claim 61 Dependent	***The Present Application***
The system of claim 14 wherein the computer program processes the images of the molecular array and extracts data from indexed features within images of the molecular array by: receiving a number of images of the molecular array produced by the scanning component; estimating initial positions of selected marker features within an image of the molecular array; calculating refined positions of the selected marker features within the image of the molecular array; using the refined positions of the selected marker features to compute an initial coordinate system for locating features of the molecular array in the number of images of the molecular array; using the initial coordinate system to locate positions of strong features within one or more images of the molecular array;	See claim 46 analysis.

refining the positions of strong features within the one or more images of the molecular array by analyzing data signal intensity values in regions of the one or more images of the molecular array that contain the strong features; using the refined positions of strong features in the one or more images of the molecular array to calculate a refined coordinate system to locate positions of weak features within the number of images of the molecular array; using the refined positions of strong features in the one or more images of the molecular array to calculate a refined coordinate system to locate positions of local background regions surrounding all strong and weak features within the number of images of the molecular array; and extracting data from strong features, and their respective local background regions, within the number of images of the molecular array using the refined positions of strong features within the number of images of the molecular array and extracting data from weak features, and their respective local background regions, within the number of images of the molecular array using locations for the weak features calculated from the refined coordinate system.	

Applicants' claims 64-65 were copied from '820 and have not been substantively examined. Applicants set out in the following table the support for each claim.

Claim 64 Independent	***Application Disclosure***	***Disclosure of Application with Feb. 1994 priority date***
A method for evaluating an orientation of a molecular array having features arranged in a pattern, the method comprising:	Page 4, lines 16-19: "In another embodiment, the invention provides a method of aligning scanned images of chips with hybridized nucleic acid sequences. A chip having attached nucleic acid sequences (probes) is synthesized, with the chip including a first pattern of nucleic acid sequences"	Figs. 6A and 6B. See page 26, lines 22-32. "Referring to FIGS. 6, the system is initialized by requesting the user to enter the name of an image file of interest. At step 601, the system retrieves the image file and prompts the user to enter the four corners of the image at

		step 602. Next, at steps 603 and 604, the system prompts the user for the number of cells located horizontally and vertically on the substrate. From the information entered by the user and the image file, the system creates a computer representation of a histogram for each cell at step 605. The histogram (at least in the form of a computer file) plots the number of pixels versus intensity."
(a) receiving an image of the molecular array produced by scanning the molecular array to determine data signals emanating from discrete positions on a surface of the molecular array;	Page 11, lines 20-21: "The image file is provided as input to analysis system 126 that incorporates the scanned image alignment techniques of the present invention" Page 11, lines 12-15: "The output of scanner 120 is an image file(s) indicating, in the case of fluorescein labeled target, the fluorescence intensity (photon counts or other related measurements, such as voltage) as a function of position on the substrate"	Much of the application discusses scanning. Figs 1A-C shows scanners. Figs. 2, 3A and B, 4 A-C and 5 show the method for generating, receiving and analyzing data from an array. See page 25, lines 29-37 "Upon completion of the conversion process, an image file representing fluorescence intensity is created and stored in memory at step 507. At step 508, the system may optionally display the image file. The intensity level of the displayed image varies from region to region according to the binding affinity of the targets to the polymer sequence therein. The brightest signals typically represent the greatest binding affinity while signals of lesser intensity represent lesser degrees of binding affinity." See page 26, lines 33- page

		27, and line 13. For example, "At step 606, the main data analysis loop is performed for each synthesis site. Analyzing the histogram for the respective synthesis site, the system calculates the total intensity level and number of pixels for the bandwidth centered around varying intensity levels."
(b) calculating an actual result of a function on pixels of the image lying in a second pattern;	(Yakhini describes the second pattern in col. 3, lines 53 - 57: "the second pattern may be a rectilinear grid of rows and columns which would lie on the rows and columns of the rectilinear grid of the array when the second pattern and the array are superimposed" Thus, Page 20, line 17: "pixel intensities on grid lines in the grid are summed", where the summing of pixel intensities is a calculation of an actual result of a function on pixels on the image that lie in the grid lines in the grid pattern.	See page 27, lines 14-20. "At step 610, an image in the form of a grid representing the substrate is displayed. Each block in the grid represents a region synthesized with a polymer sequence. The image intensity of each region will vary according to the binding affinity between the polymer sequence and targets therein. Statistical data, such as the peak and average intensity corresponding to each region are also displayed."
(c) comparing the result of step (b) with an expected result which would be obtained if the second pattern had a predetermined orientation on the array; and	Page 20, line 25 – Page 21, line 1: "Once all the positions of the grid have been analyzed, the system selects a grid position where pixel intensities (e.g. the sum calculated at step 551) are at a minimum. The system compares the summation results with an expectation of finding the lowest value, where the position of the lowest value corresponds to proper grid placement where features are not present that is predetermined to be the proper placement.	See page 27, lines 31-37. "At step 612, the system retrieves the file created during the synthesis process of the substrate being analyzed. The synthesis file contains sequence information as a function of location. The system integrates the synthesis file with the image file and sorts the data therein. Through this process, the molecular sequence of complementary probes and the intensity as a function of location is available."
(d) when the results of the	Page 21, lines 1-2: "Therefore, if	See Figs 6A and B. Steps

comparison in step (c) are outside a predetermined difference, then altering the orientation of the second pattern on the array and repeating steps (b) and (c), and repeating the foregoing as needed until the results of the comparison are within the predetermined difference.	the pixel intensities for grid lines are lower at another position, the grids is adjusted accordingly" Figure 13, elements 551, 553, 555, and Page 20, lines 22-25 illustrate the repetition of step (b): "Then, at a step 553, the system may determine if there are more positions of the grid to analyze. If there are, the position of the grid may be adjusted at a step 555. Therefore, the grid may be moved left or right by one or more pixels before the intensities are summed along the grid lines at step 551". In other words, the grid is repositioned for a number of iterations and the summation is performed at each iteration. Also, same process repeated (c) for horizontal direction (Page 21, see claim 2 description)	601 to 605 create a grid of the image data. Step 609 indicates that the method is repeated for all four corners. The expected grid (synthesis areas) is compared with the actual image in steps in 610 and 612. The image is reformatted in Step 616 and 614 and when it is matched, then displayed to the user. Also, the entire purpose for creating a grid and comparing it to the actual results is to adjust any distortion in the actual results from the image file.
Claim 65 Dependent	***Applicants' Disclosure***	
The method of claim 64 wherein: the features are arranged in a rectilinear grid and the pattern comprises a rectilinear grid of rows and columns; and step (b) comprises calculating row and column vectors by summing pixels in the rows and columns.	Figures 7A and 7B illustrate a rectilinear grid pattern having rows and columns. Page 20 lines 18-19: "the intensities of the grid in a vertical direction in the checkerboard pattern in the scanned image may be summed", and Page 21, lines 4-6: "Although the process in Fig. 13 was described for grid lines in the vertical direction, preferred embodiments also perform the same grid alignment for the horizontal direction"	

37 C.F.R. § 1.601(n) - Claims Defining The Same Patentable Invention

37 C.F.R. § 1.601(n) provides, in part, as follows:

Invention "A" is the *same patentable invention* as an invention "B" when invention "A" is the same as (35 U.S.C. 102) or is obvious (35 U.S.C. 103) in view of invention "B" assuming invention "B" is prior art with respect to invention "A".

In the context of this request for interference, a claim (*i.e.*, "Invention A") is directed to the same patentable invention as a proposed count (*i.e.*, "Invention B") when the claim is the same or is obvious in view of the proposed count, assuming the proposed count is prior art with respect to the claim. Below, Applicants explain why the identified patent and application claims define the same patentable invention as the proposed count and, therefore, should be designated as corresponding to the proposed count.

Claims 1-18 of the '196 Patent

Claims 1-18 of the '196 patent define the same patentable invention as the proposed count and, therefore, should be designated as corresponding to the proposed count for at least the following reasons:

Claim 1. Claim 1 is specifically shown in the above chart which lists the features of the claim copied from '196 versus the support in the present application. The disclosure in the present application constitutes a § 102 (e) reference because the present application issued as Fiekowsky et al., U.S. Pat. No. 6,090,555 ('555) on July 18, 2000 with a priority date of Dec. 11, 1997. Additionally, the '196 claim 1 limitations are obvious over '555 view of Shams U.S. Patent No. 6,349,144. For Example, Shams shows a method for applying a grid to a molecular array. See the flow chart in Fig. 8 where images are received (step 64), a grid is generated with and estimate of initial positions (70) and a refined or adjusted position is generated (78). Fig. 3

has a similar disclosure with receiving and storing images (16), estimating/generating (26), adjusting/refining (28) and developing an initial coordinate system (26, 28, and 30). The Shams system uses "strong features" to locate and shift grid points, see Col. 6, lines 36 to 44. Multiple iterations of the process may be performed. Col. 6, lines 65-67. Low signal areas are also identified, see Col. 9, lines 52-59. Background values are acquired during this analysis, see Col. 4, line 49, Fig. 9 (step 90), Fig. 12 (Step 104), and others. Therefore, the combination of the present application plus Shams shows that claim 1 is obvious over the count.

Claim 2. Claim 2 depends from claim 1 and claims different labels that are on the array. Again, the above chart discloses that the present application shows fluorescent and radiographic labeling techniques and this disclosure was made public in '555, if not earlier. The techniques are also shown in Stern et al U.S. Patent No. 5,631,734 issued May 20, 1997. "Fluorescent" is shown throughout the patent (even in the title) and radiographic labels are shown in the U.S. Patent No. 5,143,854, which was incorporated by reference at Col. 1. Chemiluminescent labels for arrays are shown in at least U. S. Patent No. 5,800,992 ('992) and its foreign equivalent WO 92/10588. '992 was issued September 1, 1998 and the PCT application was published around June 1991.

Claim 3. Claim 3 is dependent on claim 1 and further specifies that the images are composed of pixels having intensity values. Again, the above chart shows that these features are in the present application, which was made public in '555. These features are also shown in Stern U.S. Patent No. 5,631,734 issued May 20, 1997 in a multitude of locations, such as, Col. 8, line 58 to Col. 9, line 3. See also claim 2 of Shams.

Claim 4. Claim 4 is dependent in claim 3 and further recites arranging the array data into a grid using the corners and other information. The chart above shows how the features

of this claim are disclosed in the present application which was made public in '555. Additionally, Shams ('144) shows the selection of an image region for grid placement by defining the four corners for grid placement either by user selection or by anchor features (col. 6, line 7-14). Shams also discloses employing row and column values in the computation of a direction vector (col. 7, line 17-55), and shifting each grid paint towards regions with the highest intensity values, i.e. "peaks" (col. 6, lines 36-43).

Claim 5. Claim 5 is dependent on claim 4 and claims calculating the rows and column in various orientations. It corresponds to the count as stated in the above chart (and '555) and in view of Shams at Col. 7, line 17-55 which illustrates the calculation of direction vector. "An example calculation of the direction vector d for said bounding box size r X r is described at Col. 7, line 17-55. The bounding box is represented as a matrix P in the memory 14 with n columns and m rows, and elements p_{ij} corresponding to image intensity values at a location (i,j) in the bounding box." Thus the bounding box includes n columns and m rows where a direction vector is calculated for the bounding box. Also, Col. 6, lines 36-43: "Since it is assumed that the pixel intensity corresponding to DNA spots images 10 in the image region 18 are greater than their surrounding background 50 intensity values, the computer 34, according to the above steps, automatically shifts each grid point 24 towards local regions with the highest intensity values in subsequent iterations of said steps, wherein each grid point's location in the image frame 12 is modified". Also, Stern and Fiekowsky ('734) provides an illustrative example of histograms based upon a measure of intensity that include peaks in Fig 6A, where the histograms are analyzed to identify the distinct peaks (steps 606 and 607, and col. 16, lines 1-19.

Claim 6. Claim 6 depends from claim 3 and corresponds to the count as stated in the above chart and '555 and in view of Shams at Col. 7, line 17-55 which illustrates the calculation of direction vector. See the '196 Patent Definition of Blob analysis at Col. 10 lines

43-53 where they state "Blob analysis comprises the analysis of pixels within a region of interest encompassing the estimated position of a feature in order to first determine a threshold pixel intensity value and to then create a binary image in which all pixels having pixel intensity values greater than the threshold value are assigned the value "" 1," and all pixels having pixel intensity values in the region of interest less than the threshold value are assigned the value ""0." The coordinates of the centroid of the connected collection of pixels closest to the center of the region of interest in the binary image is then taken to be the refined pixel coordinates corresponding to the center of the feature.") The term "Binary" usually means consisting of two parts or two separate elements. Thus, for each grid point (analogous to selected marker feature) Shams discloses the combination of direction vectors d and e (two separate elements) to refine the position of the grid point. Col. 8, lines 49-51: " The computer 34 then linearly combines the direction vectors d and e to obtain a direction vector t for updating the position of the grid point 24". As illustrated in Figure 7, the grid point is substantially at the center of the feature.

Claim 7. Claim 7 is dependent on claim 6 and generally claims selecting different intensity bands for the pixels. This claim corresponds to the count and is obvious in view of U.S. Patent No. 6, 124,102 and standard statistical techniques. See for example, claims 1 and 16.

Claim 8. Claim 8 is dependent on claim 6 and also generally claims selecting different intensity bands for the pixels. This claim corresponds to the count and is obvious in view of U.S. Patent No. 6, 124,102 and standard statistical techniques. See for example, claims 1 and 16.

Claim 9. Claim 9 is dependent on claim 3 and generally claims setting a size band for a blob. This claim corresponds to the count and is obvious in view of the prior art cited for claim 6 (Shams and the present disclosure as shown in '734) and standard statistical techniques.

Claim 10. Claim 10 depends from claim 9 and is obvious over the count as stated in the above chart and '555 and in view of Shams at Col. 11, lines 41-58: where they describe displaying a plurality of DNA spot images, analyzing the background and signal intensity values and for each DNA spot relating the analyzed difference values.

Claim 11. Claim 11 is dependent on claim 3 and generally claims using linear regression analysis to produce refined features. This technique is a standard analytical tool and the claim is obvious in view of the prior art cited for claim 3 plus any standard reference for this technique.

Claim 12. Claim 12 is dependent on claim 11 and recites standard statistical methods as applied to the array analysis. This claim corresponds to the count and is obvious in view of Shams, the '555 and '734 patents, the art cited during prosecution and standard texts on the subject of statistical analysis.

Claim 13. Claim 13 is dependent on claim 1 and recites standard statistical methods as applied to the array analysis. This claim corresponds to the count and is obvious in view of Shams, the '555 and '734 patents, the art cited during prosecution and standard texts on the subject of statistical analysis.

Claim 14. Claim 14 is specifically recited in the count and therefore defines the same patentable invention as the count.

Claim 15. Claim 15 is subject to the same analysis as claim 2.

Claim 16. Claim 16 is subject to the same analysis as claim 1 above.

Claim 17. Claim 17 is dependent on claim 16 and recites multiple statistical methods to apply to features in the array. This claim corresponds to the count and is obvious in view of Shams, the '555 and '734 patents, the art cited during prosecution and standard texts on the

subject of statistical analysis.

Claim 18. Claim 18 is dependent on claim 16 and recites multiple statistical methods to apply to features in the array. This claim corresponds to the count and is obvious in view of Shams, the '555 and '734 patents, the art cited during prosecution and standard texts on the subject of statistical analysis.

Claims 1-5 of the '820 Patent

Claims 1-5 of the '820 patent define the same patentable invention as the proposed count and therefore should be designated as corresponding to the proposed count for the following reasons.

Claim 1. Claim 1 is specifically recited in the count and therefore defines the same patentable invention as the count.

Claim 2. Claim 2 is dependent on claim 1 and limits the arrangement of the features to a grid and a method for calculating vectors. It corresponds to the count or is obvious in view of it in view of this chart above and '555 and Shams. For example, Shams shows examples of features disposed in grids of rows and columns in Figs. 1, 2, 4-7 and col. 5, lines 28-47 of the specification. Shams shows summing up pixels in rows and columns in the mathematical formulae presented in col. 7, lines 23-55, where there are pixel intensity values associated with n columns and m rows used in various computations to compute an average or weighted sum of vectors described in col. 7, lines 2-16.

Claim 3. Claim 3 is dependent on claim 1 and is limited to associating a code with an array. This claim corresponds to the count or is obvious in view of the count above and Shams and others. For example, the code may be a bar code (see '820 at col. 27, line 44). Bar

codes for arrays are claimed in U.S. Patent No. 6,399,365 which has a priority date of June 8, 1994. This disclosure was published earlier as U.S. Patent No. 5,945,334 on Aug. 31, 1999 and its foreign equivalent published February 7, 1996 as EP 695 941.

Claim 4. Claim 4 is dependent on Claim 3 and is limited to obtain information from a remote location. It corresponds to the count or is of obvious in view of the count above and Shams. Layout information is obtained after synthesis and a special file is sent to the customer with the commercial array. Synthesis occurs at the manufacturing facility which is in another building from where a customer would use the array.

Claim 5. Claim 5 is dependent on claim 1 and specifies rotational orientation. Stern and Fiekowsky disclose the analysis of data to include "the steps of determining intensity as a function of substrate position" ('555 patent, at col. 2, lines 17-18), where the function of substrate position is analogous to an orientation of the data signals associated with positions on the substrate. Further, that a rotational orientation is inherently included in the range of possible orientations of 2 dimensional probe array positioned in a perpendicular plane to an incident light beam (see Figures 1A-1C with respect to substrate 230).

37 C.F.R. §1.607(a) (6) - Applicants Have Satisfied The Requirements of 35 U.S.C. § 135(b)

In accordance with 37 C.F.R. §1.607(a) (6), Applicants have complied with the requirements of 35 U.S.C. §§135(b) (1) and (2).

With regard to 35 U.S.C. § 135(b) (1), the '196 patent issued on July 8, 2003. Applicants presented claims 46-61 with an amendment filed on July 7, 2004, which is less than one year after the issue date of the '196 patent. The '820 patent issued on July 27, 2004. Applicants are

concurrently presenting an amendment copying claims 1-2 from '820 with this Request which is less than 1 year from the date of issuance. Accordingly, Applicants have satisfied the requirements of 35 U.S.C. § 135(b) (1).

With regard to 35 U.S.C. § 135(b) (2), the applications from which the '196 and '820 patents issued (USSNs. 09/896,572 and 09/659,415) were not published before issuance.

Benefit of Earlier Filed Application

For the purpose of the requested interference, Applicants are at least entitled to the benefit of the U.S. Provisional Application Ser. No. 60/069,032 ('032) filed on Dec. 11, 1997, which was later filed as USSN 09/996,731 and issued as U.S. Patent No. 6,090,555, which constitutes a constructive reduction to practice of an embodiment within the scope of the proposed count. A copy of '032 is attached. The specification of the '032 application constitutes a constructive reduction to practice of subject matter within the scope of the proposed count, as reflected in the Table below. Applicants present the benefit analysis for their claims 59 of '196 and 1 of '820, but are entitled to benefit for others claims as well. It is not necessary to show benefit for each claim in this analysis. With a concurrently filed petition, Applicants are also requesting priority to, and benefit from, another application. This application is USSN 08/195,889 filed 02/10/1994 and a copy is attached.

Claims copied from '196

Applicants' Claim 59 (recited in proposed Count)	***Support in Applicants' priority application dated Dec. 11, 1997***
A system for automated extraction of data from a molecular array having features arranged in a regular pattern, the system comprising:	Figure 3, and page 6, line 25 through page 7, line 7 describe a computerized system for forming and analyzing arrays of biological materials. Also, the analysis system is further described in page 8, lines 14-19.

	Arrays of polymers or chips are described in page 7, lines 17-19, where the chip includes a first pattern of nucleic acid sequences (page 3, lines 17-19).
a scanning component that produces images of the molecular array representing intensities of data signals emitted from discrete positions on a surface of the molecular array; a computer program that processes the images of the molecular array produced by the scanning component to index features in the images of the molecular array corresponding to molecules bound to features of the molecular array and that extracts data from the indexed features within images of the molecular array; and a computer for executing the computer program.	Figure 1, and page 8, lines 4-13 describe scanning system 120 that generates image file 124 that indicates "the fluorescence intensity (photon counts or other related measurements such as voltage) as a function of position on the substrate".
	The scanning system and scanned image are further described in page 11, lines 10-21. Exemplary computer code (i.e. a program) for processing scanned images executed by analysis system 126 is presented in the "Software Appendix", pages 18-19. Page 11, lines 18-21 describe image alignment step 263 as "the scanned image is aligned so that the pixels that correspond to each cell can be identified". Further, page 11, lines 22-27 describe step 267 that calculates the relative hybridization intensities for each cell including a calculation of the mean of the pixel values within the cell. "For example, the hybridization intensity for a cell, and therefore the relative hybridization affinity between the probe of the cell and the sample sequence, may be calculated as the mean of the pixel values within the cell" (page 15, lines 14-25). Figure 3, and page 9, line 25 through page 10, line 10 describe a computerized system for forming and analyzing arrays of biological materials. Also, the analysis system is further described in page 11, lines 20-25.

Additionally, support for claim 59 can also be found in the Stern and Fiekowsky application which is a priority application (after granting the attached Petition). Applicants concurrently file a petition with this Request to claim priority to the Stern and Fiekowsky series of applications. The priority goes back to USSN 08/195,889, which has an effective filing date of February 10, 1994. The table below shows the support for claim 59 in that earliest application.

Applicants' claim 59 (recited in proposed Count)	*Support in USSN 08/195,889*
59. A system for automated extraction of data from a molecular array having features arranged in a regular pattern, the system comprising: a scanning component that produces images of the molecular array representing intensities of data signals emitted from discrete positions on a surface of the molecular array; a computer program that processes the images of the molecular array produced by the scanning component to index features in the images of the molecular array corresponding to molecules bound to features of the molecular array and that extracts data from the indexed features within images of the molecular array; and a computer for executing the computer program.	For the scanning component, see Figs. 1-4 and the Summary of the invention as well as the remainder of the application. The Background of the Invention section and the Summary specifically describe the aspects of the scanning system. The remainder of the application goes into detail of the scanning system. For the computer component, see page 25, lines 29-37. "Upon completion of the conversion process, an image file representing fluorescence intensity is created and stored in memory at step 507. At step 508, the system may optionally display the image file. The intensity level of the displayed image varies from region to region according to the binding affinity of the targets to the polymer sequence therein. The brightest signals typically represent the greatest binding affinity while signals of lesser intensity represent lesser degrees of binding affinity." See also page 26, line 22 to page 28, line 7. Computers are an integral component of the invention as shown by the Figures and the description above. One specific passage is: "From the information entered by the user and the image file, the system creates a computer representation of a histogram for each cell at step 605. The histogram (at least in the form of a computer file) plots the number of pixels versus intensity."

	Figs. 6A and 6B specifically show a method in which an image is obtained, it is formatted into a grid, the intensities of the pixels are obtained, a grid of regions is displayed, the synthesis file is retrieved, applied to the actual image, and an analysis is performed to detect grid roam, with the potential for reformatting the image.

Claims copied from '820

Additionally, with respect to claim 64, which is also recited in the proposed count, Applicants can show a constructive reduction to practice in both the present application and the application file February 10, 1994.

Claim 64 Independent	***Disclosure in Present Application***	***Disclosure of Application with Feb. 10, 1994 priority date***
A method for evaluating an orientation of a molecular array having features arranged in a pattern, the method comprising:	Page 4, lines 16-19: "In another embodiment, the invention provides a method of aligning scanned images of chips with hybridized nucleic acid sequences. A chip having attached nucleic acid sequences (probes) is synthesized, with the chip including a first pattern of nucleic acid sequences"	Figs. 6A and 6B. See col. 15, lines 5-16. "Referring to FIGS. 6a and 6b, the system is initialized by requesting the user to enter the name of an image file of interest. At step 601, the system retrieves the image file and prompts the user to enter the four corners of the image at step 602. Next, at steps 603 and 604, the system prompts the user for the number of cells located horizontally and vertically on the substrate. From the information entered by the user and the image file, the system creates a computer representation of a histogram for each cell at step 605. The

		histogram (at least in the form of a computer file) plots the number of pixels versus intensity."
(a) receiving an image of the molecular array produced by scanning the molecular array to determine data signals emanating from discrete positions on a surface of the molecular array;	Page 11, lines 20-21: "The image file is provided as input to analysis system 126 that incorporates the scanned image alignment techniques of the present invention" Page 11, lines 12-15: "The output of scanner 120 is an image file(s) indicating, in the case of fluorescein labeled target, the fluorescence intensity (photon counts or other related measurements, such as voltage) as a function of position on the substrate"	Much of the application discusses scanning. Figs 1A-C shows scanners. Figs. 2, 3A and B, 4 A-C and 5 show the method for generating, receiving and analyzing data from an array. See col. 14, lines 43-50 "Upon completion of the conversion process, an image file representing fluorescence intensity is created and stored in memory at step 507. At step 508, the system may optionally display the image file. The intensity level of the displayed image varies from region to region according to the binding affinity of the targets to the polymer sequence therein. The brightest signals typically represent the greatest binding affinity while signals of lesser intensity represent lesser degrees of binding affinity." See col. 15, lines 17-35. For example, "At step 606, the main data analysis loop is performed for each synthesis site. Analyzing the histogram for the respective synthesis site, the system calculates the total intensity level and number of pixels for the

		bandwidth centered around varying intensity levels."
(b) calculating an actual result of a function on pixels of the image lying in a second pattern;	(Yakhini describes the second pattern in col. 3, lines 53 - 57: "the second pattern may be a rectilinear grid of rows and columns which would lie on the rows and columns of the rectilinear grid of the array when the second pattern and the array are superimposed" Thus, Page 20, line 17: "pixel intensities on grid lines in the grid are summed", where the summing of pixel intensities is a calculation of an actual result of a function on pixels on the image that lie in the grid lines in the grid pattern.	See col. 15, lines 36-42. "At step 610, an image in the form of a grid representing the substrate is displayed. Each block in the grid represents a region synthesized with a polymer sequence. The image intensity of each region will vary according to the binding affinity between the polymer sequence and targets therein. Statistical data, such as the peak and average intensity corresponding to each region are also displayed."
(c) comparing the result of step (b) with an expected result which would be obtained if the second pattern had a predetermined orientation on the array; and	Page 20, line 25 – Page 21, line 1: "Once all the positions of the grid have been analyzed, the system selects a grid position where pixel intensities (e.g. the sum calculated at step 551) are at a minimum. The system compares the summation results with an expectation of finding the lowest value, where the position of the lowest value corresponds to proper grid placement where features are not present that is predetermined to be the proper placement.	See col. 15, lines 53-59. "At step 612, the system retrieves the file created during the synthesis process of the substrate being analyzed. The synthesis file contains sequence information as a function of location. The system integrates the synthesis file with the image file and sorts the data therein. Through this process, the molecular sequence of complementary probes and the intensity as a function of location is available."
(d) when the results of the comparison in step (c) are outside a predetermined difference, then altering the orientation of the second pattern on the array and repeating steps (b) and (c), and	Page 21, lines 1-2: "Therefore, if the pixel intensities for grid lines are lower at another position, the grids is adjusted accordingly"	See Figs 6A and B. Steps 601 to 605 create a grid of the image data. Step 609 indicates that the method is repeated for all four corners. The

repeating the foregoing as needed until the results of the comparison are within the predetermined difference.	Figure 13, elements 551, 553, 555, and Page 20, lines 22-25 illustrate the repetition of step (b): "Then, at a step 553, the system may determine if there are more positions of the grid to analyze. If there are, the position of the grid may be adjusted at a step 555. Therefore, the grid may be moved left or right by one or more pixels before the intensities are summed along the grid lines at step 551". In other words, the grid is repositioned for a number of iterations and the summation is performed at each iteration. Also, same process repeated (c) for horizontal direction (Page 21, see claim 2 description)	expected grid (synthesis areas) is compared with the actual image in steps in 610 and 612. The image is reformatted in Step 616 and 614 and when it is matched, then displayed to the user. Also, the entire purpose for creating a grid and comparing it to the actual results is to adjust any distortion in the actual results from the image file.

37 C.F.R. §§ 1.601(m) and 1.608(b)

37 C.F.R. § 1.601(m) provides that a senior party in an interference is the party with the earliest effective filing date as to the count.

As noted, Applicants' present disclosure is at least entitled to the benefit of a filing date of December 11, 1997 for 60/069,032 for the claims noted above. In comparison, the earliest possible effective filing date to which the '196 and '820 patents could be accorded benefit appears to be June 6, 2000, *i.e.*, the filing date of USSN 09/589,046. Therefore, Applicants have an effective filing date of about 2 ½ years prior to the earliest possible effective filing date of the '196 and '820 patents when compared with '032. As also shown above, Applicants are also entitled to claim benefit to USSN 08/195,889, which has an effective filing date of February 10,

1994 and is 6 ½ years earlier than the priority date of '196 and '820 for various of Applicants claims.

In accordance with the provisions of 37 C.F.R. § 1.601(m), Applicants should be designated the Senior Party in the requested interference.

In view of Applicants' earlier effective filing date with respect to the '196 and '820 patents, no showing under 37 C.F.R. §1.608(b) is required.

37 C.F.R. §1.607(b) - Request to Proceed with Special Dispatch

In accordance with 37 C.F.R. §1.607(b), Applicants request that the Examiner proceed with special dispatch in declaring the requested interference.

Respectfully submitted,
AFFYMETRIX, INC.

Date: September 24, 2004

By: 
Philip L. McGarrigle
Reg. No. 31,395

Affymetrix, Inc.
Legal Department
3380 Central Expressway
Santa Clara, CA 95051
(408) 731-5000 (phone)
(408) 481-4709 (fax)

Form PTO-850 (Rev. 01-10-2001) **INTERFERENCE INITIAL MEMORANDUM** Count #____

To the Board of Patent Appeals and Interferences:

An interference is proposed involving the following __2__ parties—

PARTY	APPLICATION NO.	FILING DATE	PATENT NO., IF ANY	ISSUE DATE, IF ANY
Junior Party Yakhini et al	09/589,046	06/06/2000	6,591,196	07/08/03

If the involved is a patent, have its maintenance fees been paid? Yes ___ No ___ Not due yet __X__

Proposed priority benefit (list all intervening applications necessary for continuity):

COUNTRY	APPLICATION NO.	FILING DATE	PATENT NO., IF ANY	ISSUE DATE, IF ANY
None	None	None	None	None

The claim(s) of this party corresponding to this count:
1-18

PATENTED OR PATENTABLE PENDING CLAIMS	UNPATENTABLE PENDING CLAIMS
Patented claims 1-18	N/A

The claim(s) of this party NOT corresponding to this count:
None

PATENTED OR PATENTABLE PENDING CLAIMS	UNPATENTABLE PENDING CLAIMS
Patented Claims 1-18	N/A

PARTY	APPLICATION NO.	FILING DATE	PATENT NO., IF ANY	ISSUE DATE, IF ANY
Junior Party Yakhini et al	09/659,415	09/11/00	6,768,820	07/27/04

If the involved is a patent, have its maintenance fees been paid? Yes ___ No ___ Not due yet __X__

Proposed priority benefit (list all intervening applications necessary for continuity):

COUNTRY	APPLICATION NO.	FILING DATE	PATENT NO., IF ANY	ISSUE DATE, IF ANY
USA	09/589,046	06/06/00	6,591,196	07/08/03

The claim(s) of this party corresponding to this count:
1-5

PATENTED OR PATENTABLE PENDING CLAIMS	UNPATENTABLE PENDING CLAIMS
Patented claims 1-5	N/A

The claim(s) of this party NOT corresponding to this count:
None

PATENTED OR PATENTABLE PENDING CLAIMS	UNPATENTABLE PENDING CLAIMS
Patented Claims 1-5	N/A

PARTY	APPLICATION NO.	FILING DATE	PATENT NO., IF ANY	ISSUE DATE, IF ANY
Senior Party Fiszkowsky et al.	10/648,819	08/25/2003	N/A	N/A

If the involved is a patent, have its maintenance fees been paid? Yes___ No___ Not due yet N/A

Proposed priority benefit (list all intervening applications necessary for continuity):

COUNTRY	APPLICATION NO.	FILING DATE	PATENT NO., IF ANY	ISSUE DATE, IF ANY
USA	09/542,151	04/04/2000	6,611,767	08/26/2003
USA	09/996,737	12/23/1997	6,090,555	07/18/2000
USA	60/069,032	12/11/1997	N/A	N/A
USA	09/699,852	10/30/2000	6,741,344	05/25/2004
USA	08/823,824	03/25/1997	6,141,096	10/31/2000
USA	08/195,889	02/10/1994	5,631,734	05/20/1997

The claim(s) of this party corresponding to this count:
46-49, 59-61, and 64-65.

PATENTED OR PATENTABLE PENDING CLAIMS	UNPATENTABLE PENDING CLAIMS
Patentable pending claims 46-49, 59-61, and 64-65.	None

The claim(s) of this party NOT corresponding to this count:
None

PATENTED OR PATENTABLE PENDING CLAIMS	UNPATENTABLE PENDING CLAIMS
Patentable pending claims 46-49, 59-61, and 64-65.	None

(Check off each step, if applicable) *INSTRUCTIONS*

- ☐ 1. Obtain all files listed above.
- ☐ 2. Confirm that the proposed involved claims are still active and all corrections and entered amendments have been considered. The patents must not be expired for, among other things, failure to pay a maintenance fee (Check PALM screen 2970).
- ☐ 3. If one of the involved files is a published application or a patent, check for compliance with 35 U.S.C. 135(b).
- ☐ 4. Obtain a certified copy of any foreign benefit documents where necessary (37 CFR 1.55(a)).
- ☐ 5. Discuss the proposed interference with an Interference Practice Specialist in your Technology Center.

DATE	PRIMARY EXAMINER (signature)	ART UNIT	TELEPHONE NUMBER
DATE	INTERFERENCE PRACTICE SPECIALIST or TECHNOLOGY CENTER DIRECTOR (signature)		TELEPHONE NUMBER

PTO-103P
(Rev. 8-95)

PROVISIONAL APPLICATION
FILING RECEIPT

CORRECTED

UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office
ASSISTANT SECRETARY AND COMMISSIONER
OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

APPLICATION NUMBER	FILING DATE	FIL FEE REC'D	ATTORNEY DOCKET NO.	DRWGS
60/069,032	12/11/97	$200.00	AFFYP008+	13

MICHAEL J RITTER
HICKMAN BEYER & WEAVER
PO BOX 61059
PALO ALTO CA 94603

Receipt is acknowledged of this Provisional Application. This Provisional Application will not be examined for patentability. Be sure to provide the PROVISIONAL APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION when inquiring about this application. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. If an error is noted on this Filing Receipt, please write to Box Provisional Application within 10 days of receipt. Please provide a copy of the Provisional Application Filing Receipt with the changes noted thereon. This Provisional Application will automatically be abandoned twelve (12) months after its filing date and will not be subject to revival to restore it to pending status beyond a date which is after twelve (12) months from its filing date.

Applicant(s) PETER FIEKOWSKY, LOS ALTOS, CA; DAN BARTELL, SAN CARLOS, CA.

FOREIGN FILING LICENSE GRANTED 03/03/98 * SMALL ENTITY *
TITLE
SCANNED IMAGE ALIGNMENT SYSTEMS AND METHODS

BEYER & WEAVER, LLP

U.S.: ✓ FOREIGN: _____
DOCKETED: 8-13-98 BY: dt
ACTION: _____
DUE DATES: _____
ATTY: JKW DOCKET# _____
MJR

Attorney Docket No. AFFYP008+
(client file no. 3057)

PATENT APPLICATION

SCANNED IMAGE ALIGNMENT SYSTEMS AND METHODS

Inventors: **PETER FIEKOWSKY**

DAN BARTELL

Assignee: **AFFYMETRIX, INC.**
3380 Central Expressway
Santa Clara, California 95051
a Corporation of the state of California

Status: Large Entity

HICKMAN BEYER AND WEAVER, LLP
P.O. Box 61059
Palo Alto, CA 94306
Telephone (650) 493-6400

SCANNED IMAGE ALIGNMENT SYSTEMS AND METHODS

COPYRIGHT NOTICE

SOFTWARE APPENDICES

A Software Appendix of source code for an embodiment of the invention including two (2) sheets is included herewith.

BACKGROUND OF THE INVENTION

The present invention relates to the field of image processing. More specifically, the present invention relates to computer systems for aligning grids on a scanned image of a chip including hybridized nucleic acid sequences.

Devices and computer systems for forming and using arrays of materials on a chip or substrate are known. For example, PCT applications W092/10588 and 95/11995, both incorporated herein by reference for all purposes, describe techniques for sequencing or sequence checking nucleic acids and other materials. Arrays for performing these operations may be formed in arrays according to the methods of, for example, the pioneering techniques disclosed in U.S. Patent Nos. 5,445,934, 5,384,261 and 5,571,639, each incorporated herein by reference for all purposes.

According to one aspect of the techniques described therein, an array of nucleic acid probes is fabricated at known locations on a chip. A labeled nucleic acid is then brought into contact with the chip and a scanner generates an image file (also called a cell file) indicating the locations where the labeled nucleic acids are bound to the chip. Based upon the image file and identities of the probes at specific locations, it becomes possible to extract information such as the nucleotide or monomer sequence of DNA or RNA. Such systems have been used to form, for example, arrays of DNA that may be used to study and detect mutations relevant to genetic diseases, cancers, infectious diseases, HIV, and other genetic characteristics.

The VLSIPS™ technology provides methods of making very large arrays of oligonucleotide probes on very small chips. See U.S. Patent No. 5,143,854 and PCT patent publication Nos. WO 90/15070 and 92/10092, each of which is incorporated by reference for all purposes. The oligonucleotide probes on the DNA probe array are used to detect complementary nucleic acid sequences in a sample nucleic acid of interest (the "target" nucleic acid).

For sequence checking applications, the chip may be tiled for a specific target nucleic acid sequence. As an example, the chip may contain probes that are perfectly complementary to the target sequence and probes that differ from the target sequence by a single base mismatch. For de novo sequencing applications, the chip may include all the possible probes of a specific length. The probes are tiled on a chip in rows and columns of cells, where each cell includes multiple copies of a particular probe. Additionally, "blank" cells may be present on the chip which do not include any probes. As the blank cells contain no probes, labeled targets should not bind specifically to the chip in this area. Thus, a blank cell provides a measure of the background intensity.

Although a visual inspection of the scanned image file may be performed to identify the individual cells in the scanned image file. It would be desirable to utilize computer-implemented image processing techniques to align the scanned image file.

SUMMARY OF THE INVENTION

Embodiments of the present invention provide innovative techniques for aligning scanned images. A pattern is included in the scanned image so that when the image is convolved with a filter, a recognizable pattern is generated in the convolved image. The scanned image may then be aligned according to the position of the recognizable pattern in the convolved image. The filter may also act to remove or "filter out" the portions of the scanned image that do not correspond to the pattern in the scanned image. Several embodiments of the invention are described below.

In one embodiment, the invention provides a computer-implemented method of aligning scanned images. The scanned image is convolved with a filter. The scanned image includes a first pattern that the filter will convolve into a second pattern in the convolved image. The scanned image is then aligned according to the position of the second pattern in the convolved image. In a preferred embodiment, the first pattern may be a checkerboard pattern that is convolved into a grid pattern in the convolved image.

In another embodiment, the invention provides a method of aligning scanned images of chips with hybridized nucleic sequences. A chip having attached nucleic acid sequences (probes) is synthesized, with the chip including a first pattern of nucleic acid sequences. Labeled nucleic acid sequences are hybridized to nucleic acid sequences on the chip and the hybridized chip is scanned to produce a scanned image. The scanned image is convolved with a filter that will convolve the first pattern into a second pattern in the convolved image. The scanned image is then aligned according to the position of the second pattern in the convolved image. In a preferred embodiment, the first pattern may be a checkerboard pattern that is generated by control nucleic acid sequences that hybridize to alternating squares in the checkerboard pattern.

Other features and advantages of the invention will become readily apparent upon review of the following detailed description in association with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 illustrates an example of a computer system that may be utilized to execute the software of an embodiment of the invention.

Fig. 2 illustrates a system block diagram of the computer system of Fig. 1.

Fig. 3 illustrates an overall system for forming and analyzing arrays of biological materials such as DNA or RNA.

Fig. 4 is a high level flowchart of a process of synthesizing a chip.

Fig. 5 illustrates conceptually the binding of probes on chips.

Fig. 6 illustrates a flowchart of how a chip is hybridized and analyzed to produce experimental results.

Fig. 7A shows a checkerboard pattern in a scanned image and Fig. 7B shows a grid that has been aligned over the scanned image to show the individual cells on the chip.

Fig. 8 illustrates a flowchart of a process of image alignment.

Fig. 9A shows a checkerboard pattern in a scanned image and Fig. 9B shows a convolved image of Fig. 9A with a grid pattern that was generated by the checkerboard pattern.

Fig. 10 illustrates a flowchart of a process of convolving the scanned image.

Fig. 11 shows neighbor pixels that may be analyzed to produce a convolved pixel in the convolved image.

Figs. 12A-12D show how the filter may be moved over the scanned image to produce the convolved image.

Fig. 13 illustrates a flowchart of a process of refining the grid alignment over the scanned image.

Fig. 14 shows the grid lines in the scanned image that may be analyzed to refine the grid alignment.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Definitions

Kernel – The array of numbers that defines the effect of an image-processing convolution.

Overview

In the description that follows, the present invention will be described in reference to preferred embodiments that utilize VLSIPS™ technology for making very large arrays of oligonucleotide probes on chips. However, the invention is not limited to images produced in this fashion and may be advantageously applied other hybridization technologies or images in other technology areas. Therefore, the description of the embodiments that follows for purposes of illustration and not limitation.

Fig. 1 illustrates an example of a computer system that may be used to execute the software of an embodiment of the invention. Fig. 1 shows a computer system 1 that includes a display 3, screen 5, cabinet 7, keyboard 9, and mouse 11. Mouse 11 may have one or more buttons for interacting with a graphical user interface. Cabinet 7 houses a CD-ROM drive 13, system memory and a hard drive (*see* Fig. 2) which may be utilized to store and retrieve software programs incorporating computer code that implements the invention, data for use with the invention, and the like. Although a CD-ROM 15 is shown as an exemplary computer readable storage medium, other computer readable storage media including floppy disk, tape, flash memory, system memory, and hard drive may be utilized. Additionally, a data signal embodied in a carrier wave (*e.g.*, in a network including the Internet) may be the computer readable storage medium.

Fig. 2 shows a system block diagram of computer system 1 used to execute the software of an embodiment of the invention. As in Fig. 1, computer system 1 includes monitor 3 and keyboard 9, and mouse 11. Computer system 1 further includes subsystems such as a central processor 51, system memory 53, fixed storage 55 (*e.g.*, hard drive),

removable storage 57 (*e.g.*, CD-ROM drive), display adapter 59, sound card 61, speakers 63, and network interface 65. Other computer systems suitable for use with the invention may include additional or fewer subsystems. For example, another computer system could include more than one processor 51 (*i.e.*, a multi-processor system) or a cache memory.

The system bus architecture of computer system 1 is represented by arrows 67. However, these arrows are illustrative of any interconnection scheme serving to link the subsystems. For example, a local bus could be utilized to connect the central processor to the system memory and display adapter. Computer system 1 shown in Fig. 2 is but an example of a computer system suitable for use with the invention. Other computer architectures having different configurations of subsystems may also be utilized.

The present invention provides methods of aligning scanned images or image files of hybridized chips including nucleic acid probes. In a representative embodiment, the scanned image files include fluorescence data from a biological array, but the files may also represent other data such as radioactive intensity, light scattering, refractive index, conductivity, electroluminescence, or large molecule detection data. Therefore, the present invention is not limited to analyzing fluorescence measurements of hybridization but may be readily utilized to analyze other measurements of hybridization.

For purposes of illustration, the present invention is described as being part of a computer system that designs a chip mask, synthesizes the probes on the chip, labels the nucleic acids, and scans the hybridized nucleic acid probes. Such a system is fully described in U.S. Patent No. 5,571,639 that has been incorporated by reference for all purposes. However, the present invention may be used separately from the overall system for analyzing data generated by such systems.

Fig. 3 illustrates a computerized system for forming and analyzing arrays of biological materials such as RNA or DNA. A computer 100 is used to design arrays of biological polymers such as RNA and DNA. The computer 100 may be, for example, an appropriately programmed Sun Workstation or personal computer or workstation, such as an IBM PC equivalent, including appropriate memory and a CPU as shown in Figs. 1 and

2. The computer system 100 obtains inputs from a user regarding characteristics of a gene of interest, and other inputs regarding the desired features of the array. Optionally, the computer system may obtain information regarding a specific genetic sequence of interest from an external or internal database 102 such as GenBank. The output of the computer system 100 is a set of chip design computer files 104 in the form of, for example, a switch matrix, as described in PCT application WO 92/10092, and other associated computer files.

The chip design files are provided to a system 106 that designs the lithographic masks used in the fabrication of arrays of molecules such as DNA. The system or process 106 may include the hardware necessary to manufacture masks 110 and also the necessary computer hardware and software 108 necessary to lay the mask patterns out on the mask in an efficient manner. As with the other features in Fig. 3, such equipment may or may not be located at the same physical site but is shown together for ease of illustration in Fig. 3. The system 106 generates masks 110 or other synthesis patterns such as chrome-on-glass masks for use in the fabrication of polymer arrays.

The masks 110, as well as selected information relating to the design of the chips from system 100, are used in a synthesis system 112. Synthesis system 112 includes the necessary hardware and software used to fabricate arrays of polymers on a substrate or chip 114. For example, synthesizer 112 includes a light source 116 and a chemical flow cell 118 on which the substrate or chip 114 is placed. Mask 110 is placed between the light source and the substrate/chip, and the two are translated relative to each other at appropriate times for deprotection of selected regions of the chip. Selected chemical regents are directed through flow cell 118 for coupling to deprotected regions, as well as for washing and other operations. All operations are preferably directed by an appropriately programmed computer 119, which may or may not be the same computer as the computer(s) used in mask design and mask making.

The substrates fabricated by synthesis system 112 are optionally diced into smaller chips and exposed to marked targets. The targets may or may not be complementary to one or more of the molecules on the substrate. The targets are marked with a label such as

a fluorescein label (indicated by an asterisk in Fig. 3) and place in scanning system 120. Scanning system 120 again operates under the direction of an appropriately programmed digital computer 122, which also may or may not be the same computer as the computers used in synthesis, mask making, and mask design. The scanner 120 includes a detection device 124 such as a confocal microscope or CCD (charge-coupled device) that is used to detect the location where labeled target (*) has bound to the substrate. The output of scanner 120 is an image file(s) 124 indicating, in the case of fluorescein labeled target, the fluorescence intensity (photon counts or other related measurements, such as voltage) as a function of position on the substrate. Since higher photon counts will be observed where the labeled target has bound more strongly to the array of polymers (*e.g.*, DNA probes on the substrate), and since the monomer sequence of the polymers on the substrate is known as a function of position, it becomes possible to determine the sequence(s) of polymer(s) on the substrate that are complementary to the target.

The image file 124 is provided as input to an analysis system 126 that incorporates the scanned image alignment techniques of the present invention. Again, the analysis system may be any one of a wide variety of computer system(s), but in a preferred embodiment the analysis system is based on a Sun Workstation or equivalent. The analysis system may analyze the image file(s) to generate appropriate output 128, such as the identity of specific mutations in a target such as DNA or RNA.

Fig. 4 is a high level flowchart of a process of synthesizing a chip. At a step 201, the desired chip characteristics are input to the chip synthesis system. The chip characteristics may include (such as sequence checking systems) the genetic sequence(s) or targets that would be of interest. The sequences of interest may, for example, identify a virus, microorganism or individual. Additionally, the sequence of interest may provide information about genetic diseases, cancers or infectious diseases. Sequence selection may be provided via manual input of text files or may be from external sources such as GenBank. In a preferred embodiment that performs de novo sequencing of target nucleic acids, this steps is not necessary as the chip includes all the possible n-mer probes (where n represents the length of the nucleic acid probe).

For de novo sequencing, a chip may be synthesized to include cells containing all the possible probes of a specific length. For example, a chip may be synthesized that includes all the possible 8-mer DNA probes. Such a chip would have 65,536 cells (4*4*4*4*4*4*4*4), with each cell corresponding to a particular probe. A chip may also include other probes including all the probes of other lengths.

At a step 203 the system determines which probes would be desirable on the chip, and provides an appropriate "layout" on the chip for the probes. The layout implements desired characteristics such as an arrangement on the chip that permits "reading" of genetic sequence and/or minimization of edge effects, ease of synthesis, and the like.

The masks for the chip synthesis are designed at a step 205. The masks are designed according to the desired chip characteristics and layout. At a step 207, the system synthesizes the DNA or other polymer chips. Software controls, among other things, the relative translation of the substrate and mask, the flow of the desired reagents through a flow cell, the synthesis temperature of the flow cell, and other parameters.

Fig. 5 illustrates the binding of a particular target DNA to an array of DNA probes 114. As shown in this simple example, the following probes are formed in the array:

3'-AGAACGT

AGACCGT

AGAGCGT

AGATCGT

•

•

•

As shown, when the fluorescein-labeled (or otherwise marked) target 5'-TCTTGCA is exposed to the array, it is complementary only to the probe 3'-AGAACGT, and fluorescein will be primarily found on the surface of the chip where 3'-AGAACGT is located. The chip contains cells that include multiple copies of a particular probe. Thus,

the image file will contain fluorescence intensities, one for each probe (or cell). By analyzing the fluorescence intensities associated with a specific probe, it becomes possible to extract sequence information from such arrays using the methods of the invention disclosed herein.

For ease of reference, one may call bases by assigning the bases the following codes:

Code	Group	Meaning
A	A	Adenine
C	C	Cytosine
G	G	Guanine
T	T(U)	Thymine (Uracil)
M	A or C	aMino
R	A or G	puRine
W	A or T(U)	Weak interaction (2 H bonds)
Y	C or T(U)	pYrimidine
S	C or G	Strong interaction (3 H bonds)
K	G or T(U)	Keto
V	A, C or G	not T(U)
H	A, C or T(U)	not G
D	A, G or T(U)	not C
B	C, G or T(U)	not A
N	A, C, G, or T(U)	Insufficient intensity to call
X	A, C, G, or T(U)	Insufficient discrimination to call

Most of the codes conform to the IUPAC standard. However, code N has been redefined and code X has been added.

Scanned Image Alignment

Before the scanned image alignment of the invention are discussed, it may be helpful to provide an overview of the overall process in one embodiment. Fig. 6 illustrates a flowchart of a process of how a chip is hybridized and analyzed to produce experimental results. A chip 251 having attached nucleic acid sequences (or probes) is combined with a sample nucleic acid sequence (*e.g.*, labeled fragments of the sample) and reagents in a hybridization step 255. The hybridization step produces a hybridized chip 257.

The hybridized chip is scanned at a step 259. For example, the hybridized chip may be laser scanned to detect where fluorescein-labeled sample fragments have hybridized to the chip. Numerous techniques may be utilized to label the sample fragments and the scanning process will typically be performed according to the type of label utilized. The scanning step produces a digital image of the chip.

In preferred embodiments, the scanned image of the chip includes varying fluorescent intensities that correspond to the hybridization intensity or affinity of the sample to the probes in a cell. In order to achieve more accurate results, it is beneficial to identify the pixels that belong to each cell on the chip. At an image alignment step 263, the scanned image is aligned so that the pixels that correspond to each cell can be identified. Optionally, the image alignment step includes the alignment of a grid over the scanned image (*see* Fig. 7B).

At a step 267, the analysis system analyzes the scanned image to calculate the relative hybridization intensities for each cell of interest on the chip. For example, the hybridization intensity for a cell, and therefore the relative hybridization affinity between the probe of the cell and the sample sequence, may be calculated as the mean of the pixel values within the cell. The pixel values may correspond to photon counts from the labeled hybridized sample fragments.

The cell intensities may be stored as a cell intensity file 269. At an analysis step 271, the analysis system may analyze the cell intensity file and chip characteristics to

generate results 273. The chip characteristics may be utilized to identify the probes that have been synthesized at each cell on the chip. By analyzing both the sequence of the probes and their hybridization intensities from the cell intensity file, the system is able to extract sequence information such as the location of mutations, deletions or insertions, or the sequence of the sample nucleic acid. Accordingly, the results may include sequence information, graphs of the hybridization intensities of probe(s), graphs of the differences between sequences, and the like. See U.S. Patent Application No. 08/327,525, which is hereby incorporated by reference for all purposes.

In order to align the scanned image, the invention provides a pattern in the scanned image that will be convolved into a recognizable pattern. In preferred embodiments, the pattern in the scanned image is a checkerboard pattern that is generated by synthesizing alternating cells that include probes that are complementary to a control nucleic acid sequence. The control nucleic acid sequence may be a known sequence that is labeled and hybridized to the chip for the purpose of aligning the scanned image. Additionally, the brightness of the cells complementary to the control nucleic acid sequence may be utilized as a baseline or for comparison to other intensities.

As an example, Fig. 7A shows a checkerboard pattern in a hybridized chip. A scanned image 301 of a hybridized chip includes an active area 303 where the probes were synthesized. At the corner of the active area is a pattern 305 that is a checkerboard pattern. Typically, the pattern appears at each corner of the active area of the scanned image. Although the pattern is shown as being a checkerboard pattern, in other embodiments the pattern is a circle, square, plus sign, or any other pattern.

With regard to Fig. 6, it was stated that a grid may optionally be placed over the scanned image to show or delineate the individual cells of the chip. Fig. 7B shows a grid that has been aligned over the scanned image of Fig. 7A to show the individual cells of the chip. As shown, a grid 307 has been placed over active area 303 of hybridized chip 301.

Fig. 8 illustrates a flowchart of a process of image alignment. The flowchart shows detail for step 263 of Fig. 6. At a step 351, the scanned image is convolved with a filter. The filter is typically a software filter that convolves the scanned image into a

convolved image. When the scanned image is convolved, a pattern in the scanned image is convolved into a recognizable pattern. The position of the recognizable pattern in the convolved image may be utilized to align the scanned image, such as by placing a grid over the image.

At a step 353, the convolved image is searched for bright areas. When the scanned image is convolved, the pattern(s) in the scanned image will be convolved into a recognizable pattern or patterns of bright areas. Accordingly, once bright areas are identified in the convolved image, the system confirms that the bright areas are in the expected recognizable pattern (*e.g.*, a grid pattern) at a step 355.

In order to better understand what is meant by the different patterns, Fig. 9A shows a checkerboard pattern 401 in a scanned image 403. Fig. 9B shows a recognizable pattern 451 in convolved image 453. The convolved image was generated from the scanned image of Fig. 9A. As shown, recognizable pattern 41 in this embodiment is a grid pattern that was generated by the checkerboard pattern when it was convolved with a filter. Additionally, it should be noted that the filter acted to remove the other cell intensities so that the convolved image only includes the recognizable pattern. By removing cells that are not part of the pattern in the scanned image, it is easier to align the scanned image.

Fig. 10 illustrates a flowchart of a process of convolving the scanned image. The flowchart illustrates a process that may be performed at step 351 of Fig. 8. At a step 501, a pixel is selected. For simplicity, we will assume that the process selects pixels of the scanned image from left to right and top to bottom. Of course, the order that the pixels are analyzed may be varied.

Once a pixel selected, neighbor pixels may then be selected at a step 503. By neighbor pixels, it is meant pixels that the pixels are near, but not necessarily adjacent to a pixel. For example, Fig. 11 shows neighbor pixels that may be analyzed to produce a convolved pixel in a convolved image. As shown in Fig. 11, there are 9 pixels labeled 1-9. In a preferred embodiment, pixel 1 is the pixel retrieved at step 501 and the neighbor

pixels retrieved at step 503 are pixels 2-9. Of course, any number or location of different neighbor pixels may be utilized.

At a step 505, the average of the odd pixels and the average of the even pixels is determined. Referring again to Fig. 11, the intensities of pixels 1, 3, 5, 7, and 9 may be averaged to produce the average of the odd pixels (AVG_O). Similarly, the intensities of pixels 2, 4, 6, and 8 may be averaged to produce the average of the odd pixels (AVG_E).

Pixel 1 is convolved into a convolved pixel in a convolved image by determining if the average of the odd pixels is greater than the average of the even pixels at a step 507. If the average of the odd pixels is greater, the convolved pixel is set equal to the intensity of the minimum of the odd pixels minus the intensity of the maximum of the even pixels at a step 509. Otherwise, the convolved pixel is set equal to the intensity of the minimum of the even pixels minus the intensity of the maximum of the odd pixels at a step 511.

Conceptually, the neighbor pixels may be thought of as being filtered, such as by a software filter in preferred embodiments. With the filter, the system is searching for a checkerboard pattern where all the odd pixels are either darker or lighter than the even pixels. Accordingly, averages of the odd and even pixels are calculated at step 505. Step 507 acts to determine if the pixels likely reflect a checkerboard pattern where the odd pixels, and therefore squares, are light (*e.g.*, high intensity) or dark (*e.g.*, low intensity). If the odd pixels likely reflect a checkerboard pattern where the odd pixels are light, step 509 sets the convolved pixel to the difference between selected odd and even pixels, where the selected odd pixel is the minimum of the odd pixels and the selected even pixel is the maximum of the even pixels. Step 511 is similar but reversed.

Therefore, at step 509, if all the odd pixels are much brighter than all the even pixels, the difference will be a larger value. Hence, the convolved pixel will be relatively bright (*e.g.*, high intensity). The convolved pixel will also be relatively bright if all the even pixels are much brighter than all the odd pixels at step 511. However, if the difference at step 509 or 511 is very small (or negative), the convolved pixel will be set to a relatively dark intensity. Therefore, if the filter finds a checkerboard pattern, the

convolved pixel will be bright and if the filter finds a relatively random pattern, the convolved pixel will be dark (thus, filtering out "noise" that is not the desired pattern).

The recognizable pattern in Fig. 9B, which is a grid pattern, was generated by the software filter of Fig. 10. In order to better see how the recognizable pattern was generated, Figs. 12A-D show how the filter may be moved over the checkerboard to produce a grid pattern in the convolved image. As the filter is over the pattern in the scanned image shown in a square 530 in Fig. 12A, a bright square will be generated in the convolved image since a checkerboard pattern will be found. Similarly, a bright square will be generated in the convolved image when the filter is over the pattern in square 530 of Fig. 12B. Of course, the checkerboard patterns in square 530 of Figs. 12A and 12B are reversed, but both will produce a bright square in the convolved image as described above in reference to Fig. 10. Figs. 12C and 12D will also produce two bright squares. Therefore, a 2x2 bright square grid pattern is generated as shown in Fig. 9B.

Additionally, as the software filter of Fig. 10 acts to filter out signals that are not the desired pattern, the recognizable pattern (*e.g.*, a grid pattern) is easier to identify. The recognizable patterns in the convolved image are utilized to align the scanned image. Returning now to Fig. 10, after a selected pixel is convolved into a convolved pixel by the filter, it is determined if there is another pixel to process in the scanned image at a step 513.

The following shows how well an embodiment of the invention aligned scanned images of hybridized chips:

	Old method	With filter convolution
Perfect alignment	0%	4%
1 pixel off	8%	96%
2 or more pixels off	20%	0%
1 or more cells off	12%	0%
unable to align	60%	0%

As shown, an embodiment of the invention was able to dramatically increase the accuracy of scanned image alignment.

Refined Grid Alignment

In preferred embodiments, refined image alignment may be performed to further increase the accuracy of the scanned image alignment. Fig. 13 illustrates a flowchart of a process of refining grid alignment over a scanned image. Thus, for example, once the above-described process has been performed to align the scanned image, the process in Fig. 13 may be utilized to refine the alignment.

At a step 551, pixel intensities on grid lines in the grid are summed. For example, the intensities of the grid in a vertical direction in the checkerboard pattern in the scanned image may be summed. Fig. 14 shows the grid lines in the scanned image that may be analyzed to refine the grid alignment. As shown, the pixel intensities of vertical lines 601 of a checkerboard pattern 603 may be summed and stored.

Then, at a step 553, the system may determine if there are more positions of the grid to analyze. If there are, the position of the grid may be adjusted at a step 555. Therefore, the grid may be moved left and right by one or more pixels before the intensities are summed along grid lines at step 551. Once all the positions of the grid have been analyzed, the system selects a grid position where pixel intensities (*e.g.*, the sum calculated at step 551) are at a minimum. Therefore, if the pixel intensities for grid lines are lower at another position, the grid is adjusted accordingly. This refinement will work well if the cells are typically separated by a darker area or line.

Although the process in Fig. 13 was described for grid lines in the vertical direction, preferred embodiments also perform the same grid alignment for the horizontal direction. The distance that the grid is able to be moved for refinement may be limited. For example, the grid may be limited to movement of one-third a cell size.

The following shows how well an embodiment of the invention aligned scanned images of hybridized chips utilizing the refined grid alignment:

	Old method	With refined grid alignment
Perfect alignment	0%	64%
1 pixel off	8%	36%

2 or more pixels off	20%	0%
1 or more cells off	12%	0%
unable to align	60%	0%

As shown, an embodiment of the invention was able to dramatically increase the accuracy of scanned image alignment. Furthermore, refining grid alignment increased the percentage of scanned images that were perfectly aligned with the invention from 4% to 64%. Therefore, performing a refinement of grid alignment can significantly increase the accuracy of the grid alignment.

Conclusion

While the above is a complete description of preferred embodiments of the invention, various alternatives, modifications, and equivalents may be used. It should be evident that the invention is equally applicable by making appropriate modifications to the embodiments described above. For example, the invention has been described in reference to a checkerboard pattern in the scanned image. However, the invention is not limited to any one pattern and may be advantageously applied to other patterns including those described herein. Therefore, the above description should not be taken as limiting the scope of the invention that is defined by the metes and bounds of the appended claims along with their full scope of equivalents.

Software listing of the algorithm:

```
////////////////////////////////////////////////////////////////////////
// CheckerFilt
// purpose
//    perform a checker-board kernel filter on the image.
// input
//    cellWidth, cellHeight, size of the cell
//    *img, the #of rows and columns in the image and the image data
// output
//    *img, the image is filtered in place

void CheckerFilt(int cellWidth, int cellHeight, IMAGE *img)
{
   int row,col,rowBegin,nRows,nCols,colBegin,rowEndFilter,colEndFilter,imgOffset;
   int oddAvg,evenAvg,oddMin,oddMax,evenMin,evenMax;
   int temp;
   PIX_T *e1=NULL,*e2=NULL,*e3=NULL,*e4=NULL,*e5=NULL,*e6=NULL,*e7=NULL,*e8=NULL,*e9=NULL;
   //
   // Determine the range of rows and columns to filter
   rowBegin=0;
   colBegin=0;
   nRows=img->rows;
   nCols=img->cols;
   rowEndFilter=nRows-1-2*cellHeight;
   colEndFilter=nCols-1-2*cellWidth;

   //For each row
   for (row=rowBegin;row<=rowEndFilter;row++)
   {
      //Initialize the filter's pointers
      // e1  e2  e3
      // e4  e5  e6
      // e7  e8  e9
      //
      Set3x3Pointers(img, row, cellWidth, cellHeight, &e1,&e2,&e3,&e4,&e5,&e6,&e7,&e8,&e9);

      // walk the row, doing the filter
      for(col=colBegin;col<=colEndFilter;col++)
      {
         // Avg1 = Average pixels 1, 3, 5, 7, 9
         // Avg2 = Average pixels 2, 4, 6, 8
         oddAvg = (e1[col] + e3[col] + e5[col] + e7[col] + e9[col])/5;
         evenAvg = (e2[col] + e4[col] + e6[col] + e8[col])/4;

         // If  avgOdd > avgEven
         //   Then the area is bright and
         //    NewPixel = min(v1,v3,v5,v7,v9) - max(v2,v4,v6,v8)
         //   Else the area is dark and
         //    NewPixel = min(v2,v4,v6,v8) - max(v1,v3,v5,v7,v9)
         //
         if (oddAvg > evenAvg)
         {
            oddMin=MIN(e1[col],MIN(e3[col], MIN(e5[col], MIN(e7[col], e9[col]))));
            evenMax = MAX(e2[col],MAX(e4[col], MAX(e6[col], e8[col])));
            e1[col]= MAX(0,oddMin-evenMax);
            temp=e1[col];
            if(temp >0)
             temp=e1[col];
         }
```

```
        else
        {
           evenMin = MIN(e2[col],MIN(e4[col], MIN(e6[col], e8[col])));
           oddMax=MAX(e1[col],MAX(e3[col], MAX(e5[col], MAX(e7[col], e9[col]))));
           e1[col] =  MAX(0,evenMin - oddMax);
           temp=e1[col];
           if(temp >0)
            temp=e1[col];

        }
      }
   }
   //Set the border pixels, which are not filtered, to zero.
   for(row=0;row<nRows;row++)
   {
      imgOffset=row* (img->cols);
      e1=img->image+imgOffset;
      if(row<rowEndFilter)
         colBegin=colEndFilter;
      else
         colBegin=0;
      for(col=colBegin;col<nCols;col++)
         e1[col]=0;
   }
   return;
}
////////////////////////////////////////////////////////////////////////////
// Set3x3Pointers
// purpose
//    intialize pointers that will be used when walking the kernel along
//    a row of image data.
// input
//    *img: image struct contains number of rows and columns in the image
//    row:  the row of the image on which we are applying the kernel
//    cellWidth, cellHeight: size of the cell which implies the size of the kernel
//
// output
//    e1..e9:  pointers to the 9 pixels that will be used for kernel calculations
//
void Set3x3Pointers(IMAGE *img,int row, int cellWidth, int cellHeight,
  PIX_T **e1,PIX_T **e2,PIX_T **e3,PIX_T **e4,PIX_T **e5,PIX_T **e6,
  PIX_T **e7,PIX_T **e8,PIX_T **e9)
{
   PIX_T *p1=NULL,*p2=NULL,*p3=NULL;
   int imgOffset;
   int cellWidthTimes2=cellWidth*2;
   int nCols=img->cols;

   imgOffset=row*(img->cols);

   p1=img->image+imgOffset;
   p2=p1+nCols*cellHeight;
   p3=p1+nCols*2*cellHeight;
   *e1 = p1; *e2 = p1+cellWidth; *e3 = p1+cellWidthTimes2; /* SET THE POINTERS FOR THE 3 ROWS */
   *e4 = p2; *e5 = p2+cellWidth; *e6 = p2+cellWidthTimes2; /* (WHOSE POINTERS ROTATE) */
   *e7 = p3; *e8 = p3+cellWidth; *e9 = p3+cellWidthTimes2;
}
```

CLAIMS

What is claimed is:

1. In a computer system, a method of aligning scanned images, comprising:

convolving a scanned image with a filter, the scanned image including a first pattern that the filter will convolve into a second pattern in a convolved image; and

aligning the scanned image according to a position of the second pattern in the convolved image.

2. The method of claim 1, wherein convolving a scanned image with a filter comprises setting a convolved pixel to a difference between a selected odd pixel and a selected even pixel of the first pattern.

3. The method of claim 2, wherein the selected odd pixel has the lowest intensity of the odd pixels and the selected even pixel has the highest intensity of the even pixels, if the average intensity of the odd pixels is greater than the average intensity of the even pixels.

4. The method of claim 2, wherein the selected odd pixel has the highest intensity of the odd pixels and the selected even pixel has the lowest intensity of the even pixels, if the average intensity of the odd pixels is not greater than the average intensity of the even pixels.

5. The method of claim 1, wherein the first pattern is a checkerboard pattern.

6. The method of claim 1, wherein the second pattern is a grid pattern.

7. The method of claim 1, wherein aligning the scanned image comprises aligning a grid over the scanned image.

8. The method of claim 7, further comprising adjusting the position of the grid to minimize a sum of the intensities of pixels along a direction in the grid.

9. The method of claim 1, wherein the scanned image includes multiple copies of the first pattern.

10. The method of claim 9, wherein the scanned image is a rectangle with a copy of the first pattern near each corner.

11. A computer program product that aligns scanned images, comprising:

computer code that convolves a scanned image with a filter, the scanned image including a first pattern that the filter will convolve into a second pattern in a convolved image;

computer code that aligns the scanned image according to a position of the second pattern in the convolved image; and

a computer readable medium that stores the computer codes.

12. A method of aligning scanned images, comprising:

synthesizing a chip having attached nucleic acid sequences, the chip including a first pattern of nucleic acid sequences;

hybridizing labeled nucleic acid sequences to nucleic acid sequences on the chip;

scanning the hybridized chip to produce a scanned image;

convolving the scanned image with a filter, the filter convolving the first pattern into a second pattern in a convolved image; and

aligning the scanned image according to a position of the second pattern in the convolved image.

13. The method of claim 12, wherein convolving the scanned image with a filter comprises setting a convolved pixel to a difference between a selected odd pixel and a selected even pixel of the first pattern.

14. The method of claim 13, wherein the selected odd pixel has the lowest intensity of the odd pixels and the selected even pixel has the highest intensity of the even pixels, if the average intensity of the odd pixels is greater than the average intensity of the even pixels.

15. The method of claim 13, wherein the selected odd pixel has the highest intensity of the odd pixels and the selected even pixel has the lowest intensity of the even pixels, if the average intensity of the odd pixels is not greater than the average intensity of the even pixels.

16. The method of claim 12, wherein the first pattern is a checkerboard pattern.

17. The method of claim 16, wherein the labeled nucleic acid sequences include control nucleic acid sequences that hybridize to alternating squares in the checkerboard pattern.

18. The method of claim 12, wherein the second pattern is a grid pattern.

19. The method of claim 12, wherein aligning the scanned image comprises aligning a grid over the scanned image.

20. The method of claim 19, further comprising adjusting the position of the grid to minimize a sum of the intensities of pixels along a direction in the grid.

21. The method of claim 12, wherein the scanned image includes multiple copies of the first pattern.

22. The method of claim 21, wherein the scanned image is a rectangle with a copy of the first pattern near each corner.

23. A computer program product that aligns scanned images, comprising:

computer code that receives as input a scanned image of a chip having attached nucleic acid sequences to which labeled nucleic acid sequences are hybridized, the chip including a first pattern of nucleic acid sequences;

computer code that convolves the scanned image with a filter, the filter convolving the first pattern into a second pattern in a convolved image;

computer code that aligns the scanned image according to a position of the second pattern in the convolved image; and

a computer readable medium that stores the computer codes.

SCANNED IMAGE ALIGNMENT SYSTEMS AND METHODS

ABSTRACT

Systems and methods for aligning scanned images are provided. A pattern is included in the scanned image so that when the image is convolved with a filter, a recognizable pattern is generated in the convolved image. The scanned image may then be aligned according to the position of the recognizable pattern in the convolved image. The filter may also act to remove the portions of the scanned image that do not correspond to the pattern in the scanned image.

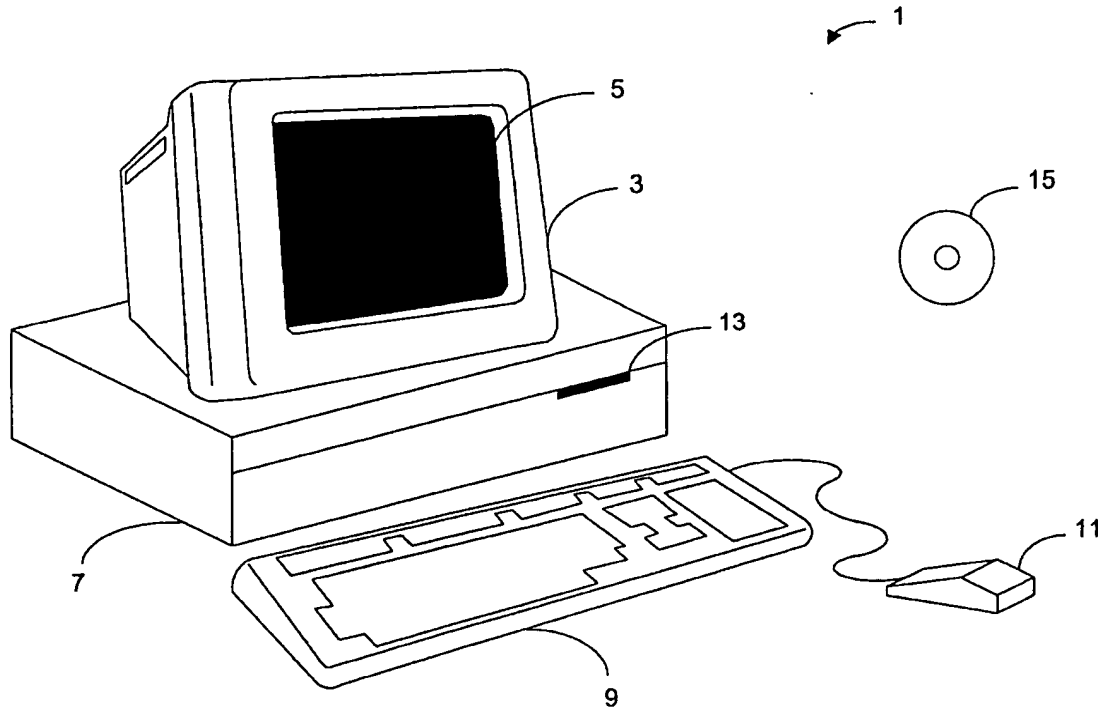

1
5
3
15
13
11
7
9

FIG. 1

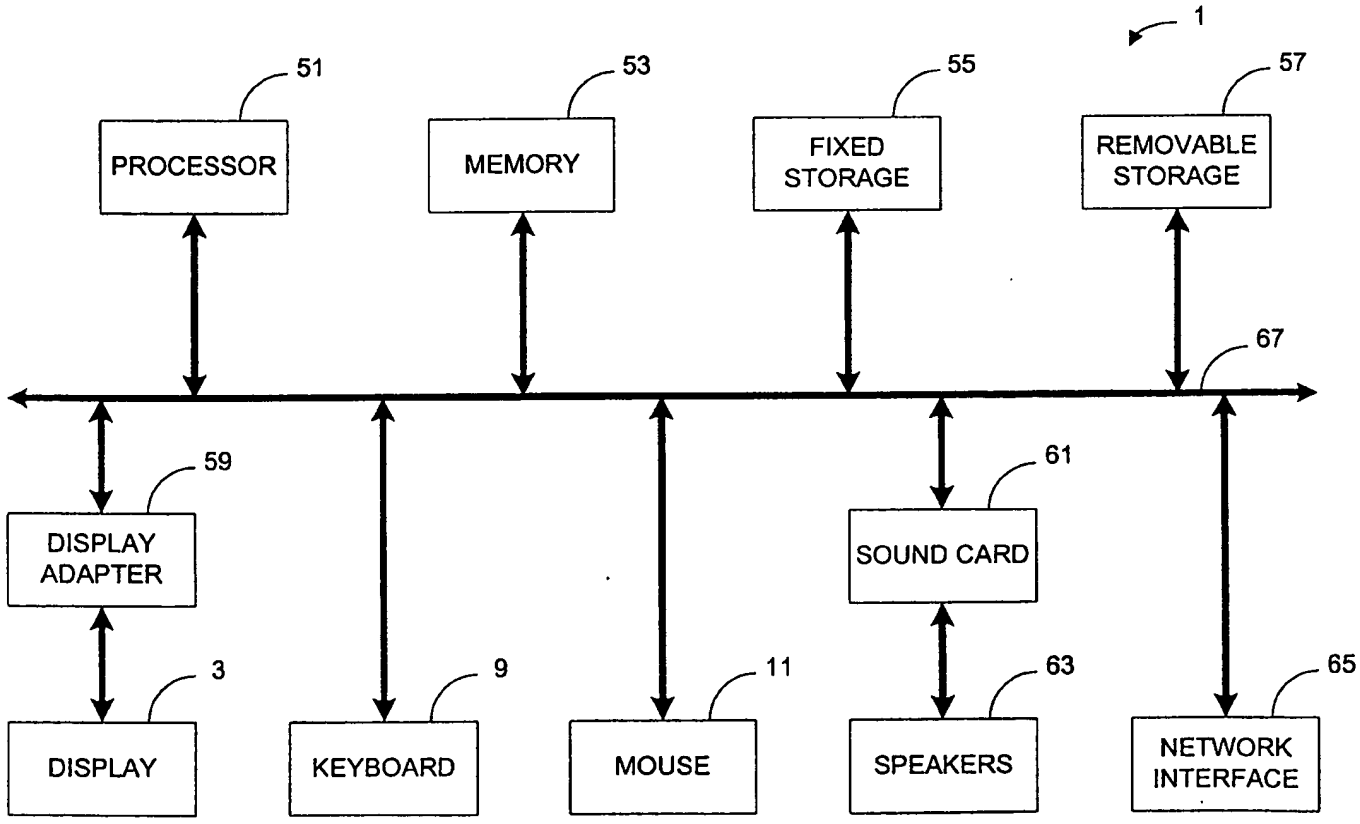

1
51
PROCESSOR
53
MEMORY
55
FIXED STORAGE
57
REMOVABLE STORAGE
67
59
DISPLAY ADAPTER
61
SOUND CARD
3
DISPLAY
9
KEYBOARD
11
MOUSE
63
SPEAKERS
65
NETWORK INTERFACE

FIG. 2

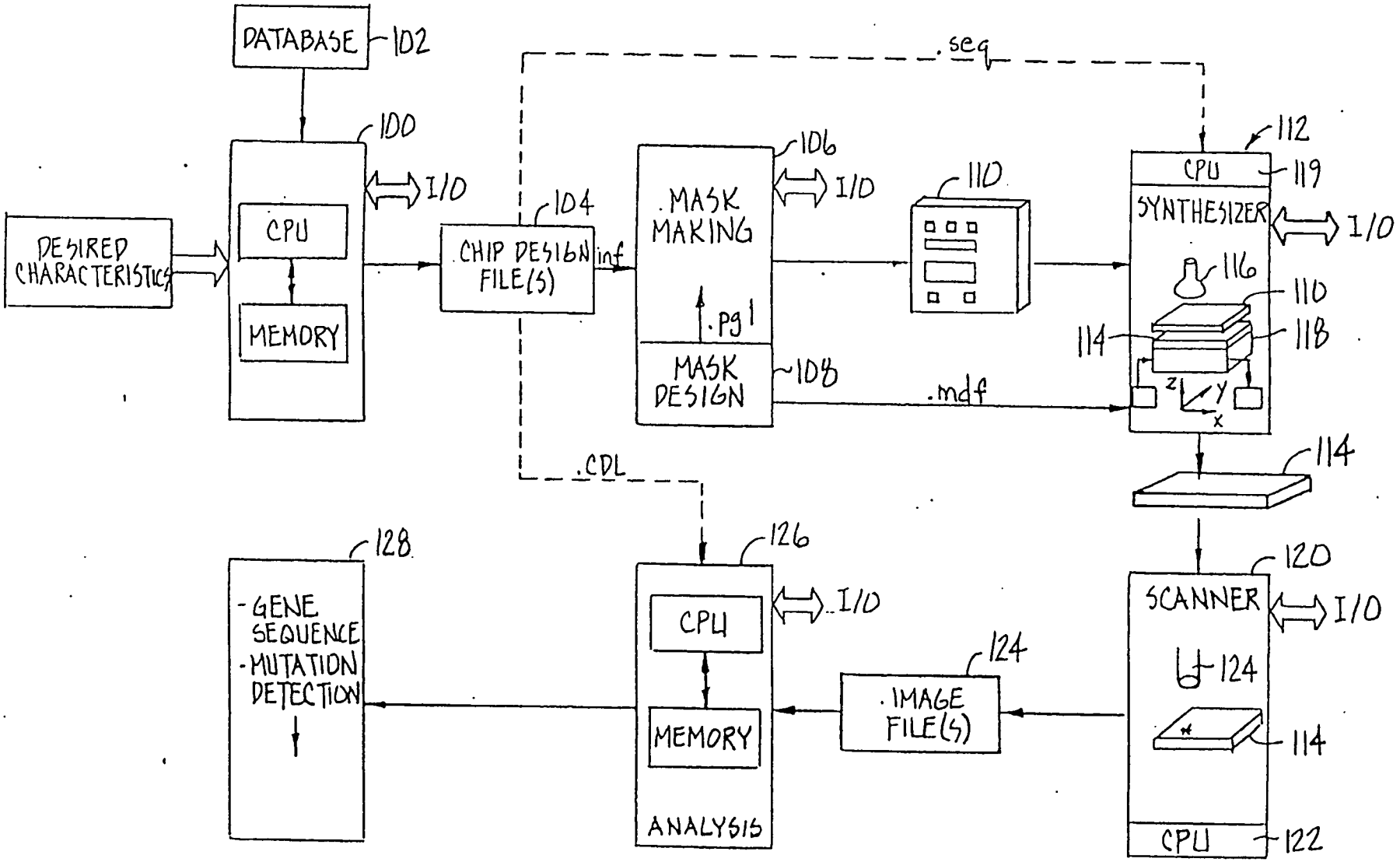

DATABASE
102
100
CPU
MEMORY
I/O
DESIRED CHARACTERISTICS
104
CHIP DESIGN FILE(S)
inf
.seq
106
MASK MAKING
I/O
.pgl
MASK DESIGN
108
.mdf
110
112
CPU
119
SYNTHESIZER
I/O
116
110
114
118
z
y
x
114
.CDL
126
CPU
I/O
MEMORY
ANALYSIS
128
- GENE SEQUENCE
- MUTATION DETECTION
124
IMAGE FILE(S)
120
SCANNER
I/O
124
114
CPU
122

FIG. 3

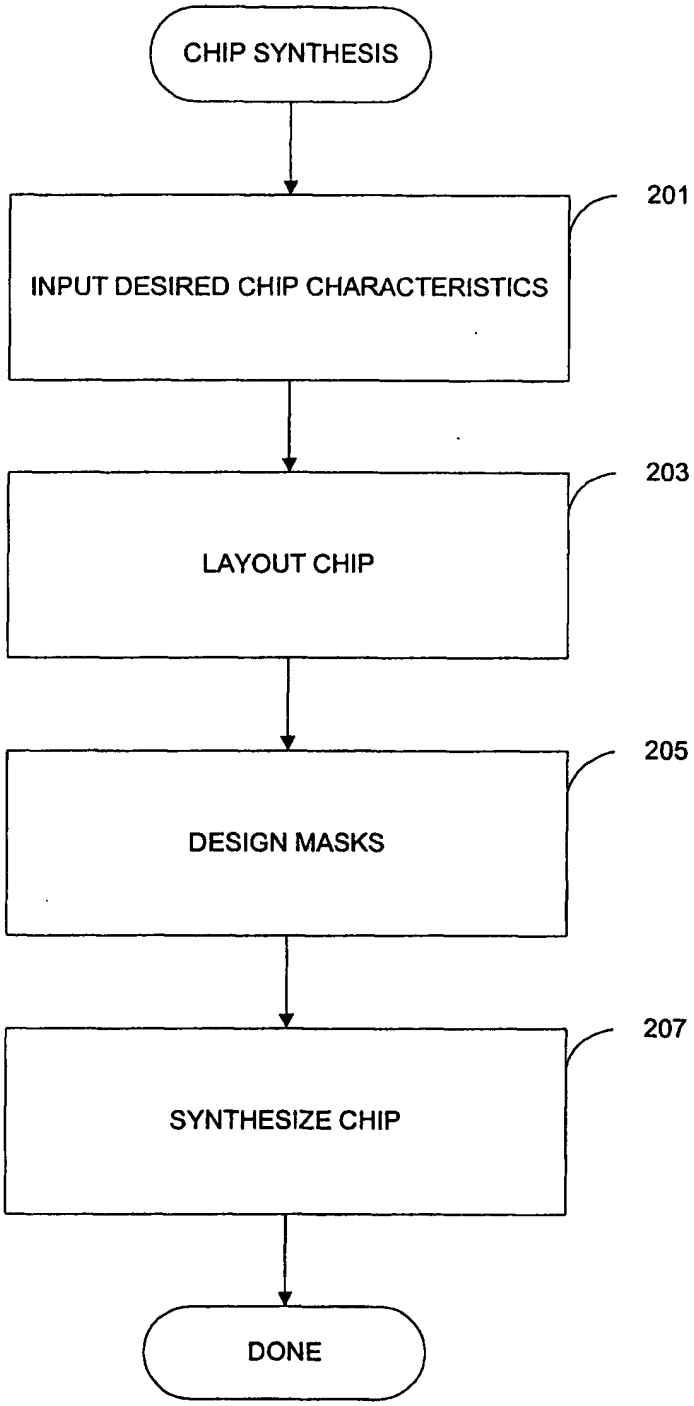

CHIP SYNTHESIS
INPUT DESIRED CHIP CHARACTERISTICS
201
LAYOUT CHIP
203
DESIGN MASKS
205
SYNTHESIZE CHIP
207
DONE

FIG. 4

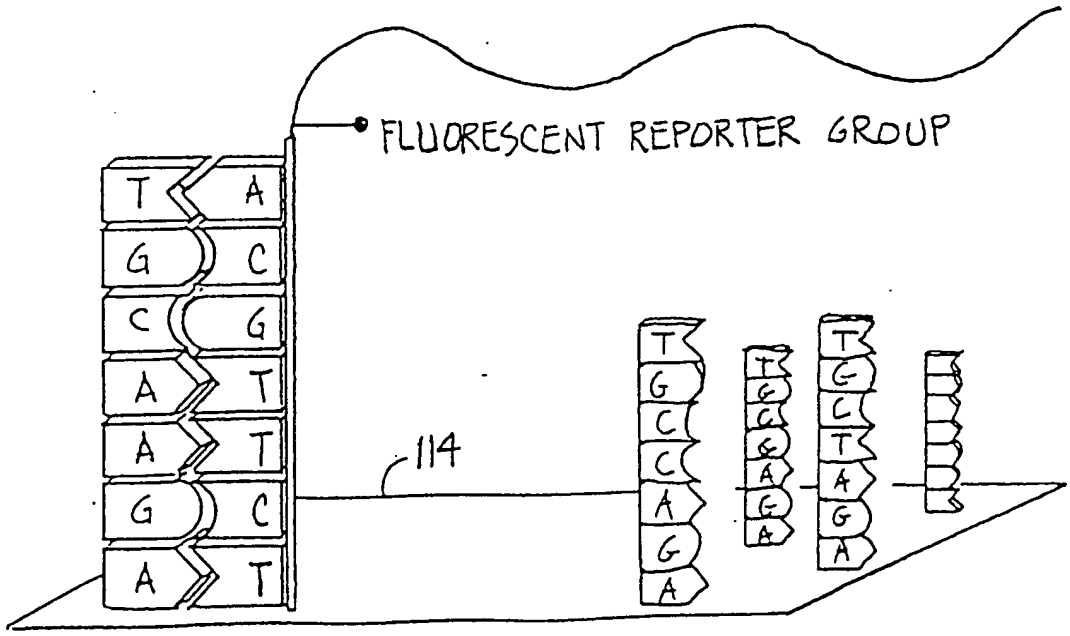


FIG. 5

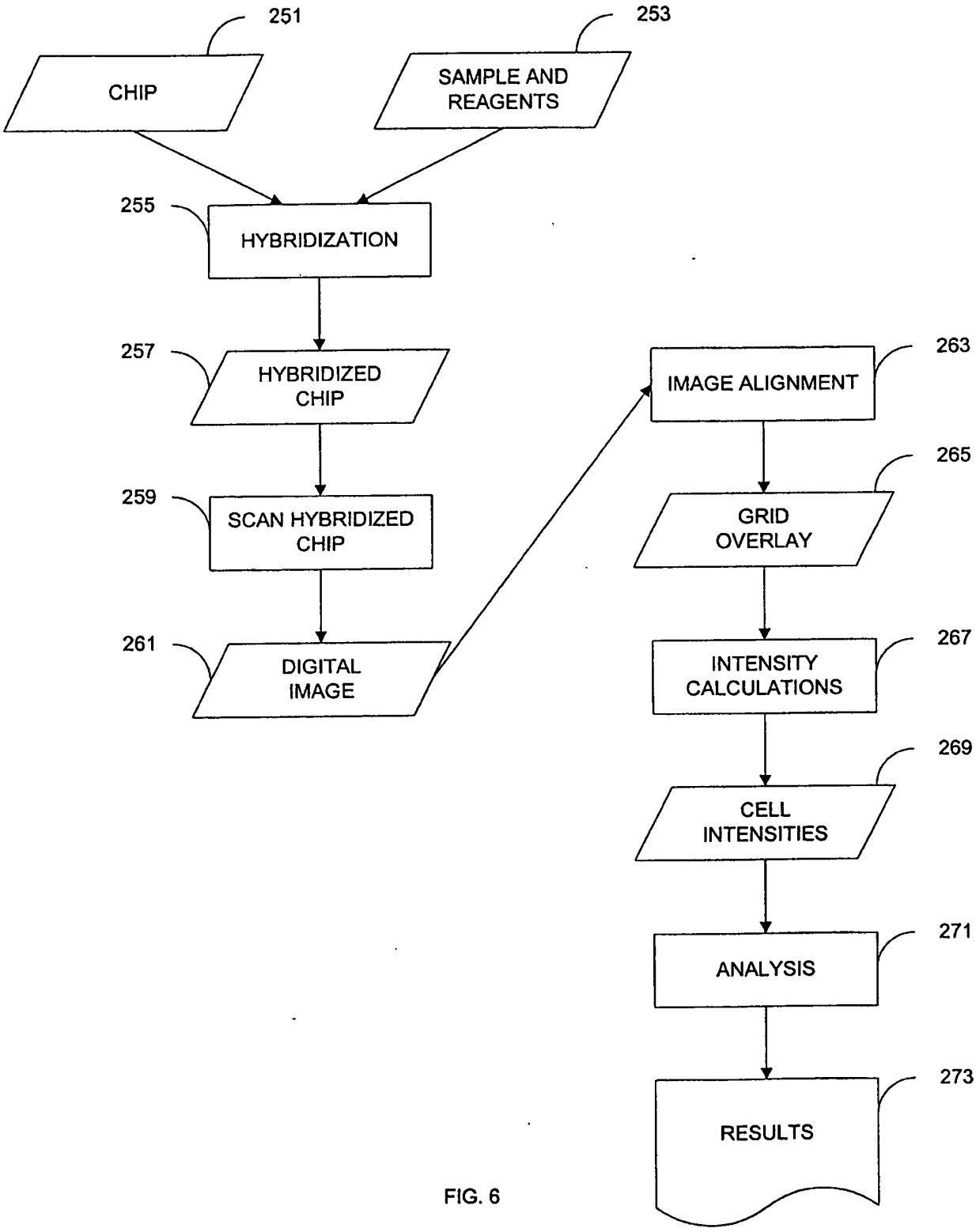

251
CHIP
253
SAMPLE AND REAGENTS
255
HYBRIDIZATION
257
HYBRIDIZED CHIP
259
SCAN HYBRIDIZED CHIP
261
DIGITAL IMAGE
263
IMAGE ALIGNMENT
265
GRID OVERLAY
267
INTENSITY CALCULATIONS
269
CELL INTENSITIES
271
ANALYSIS
273
RESULTS

FIG. 6

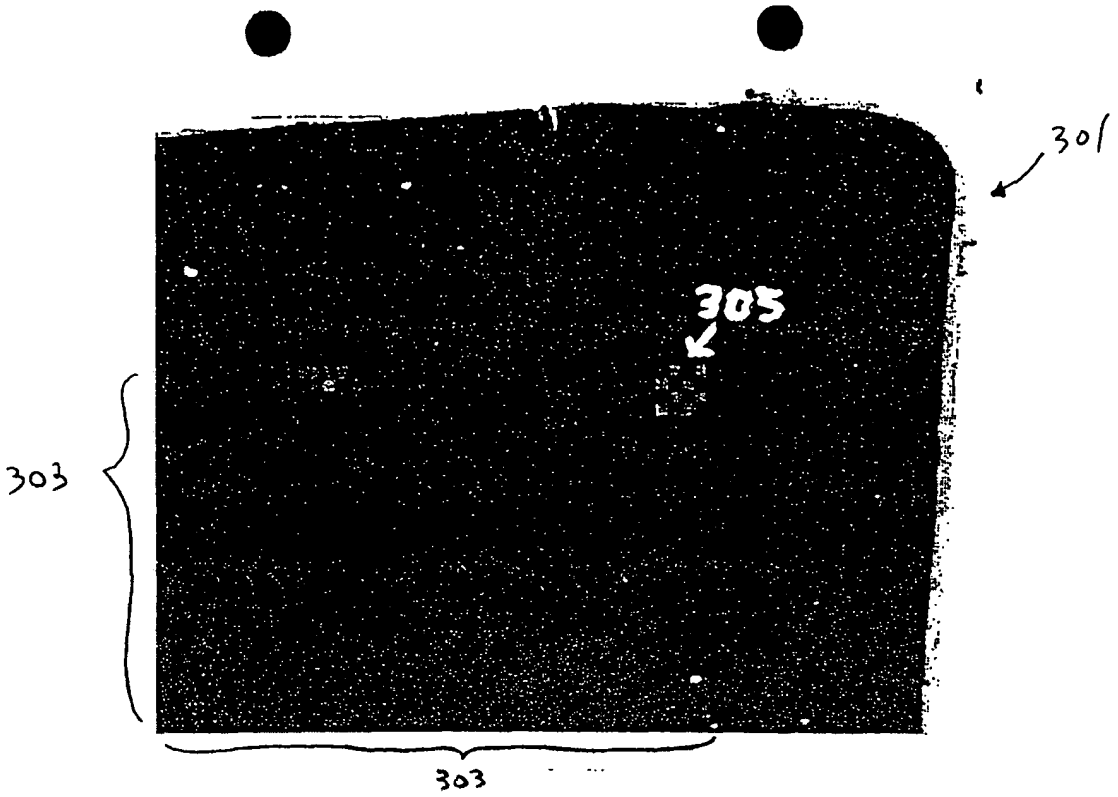


FIG. 7A

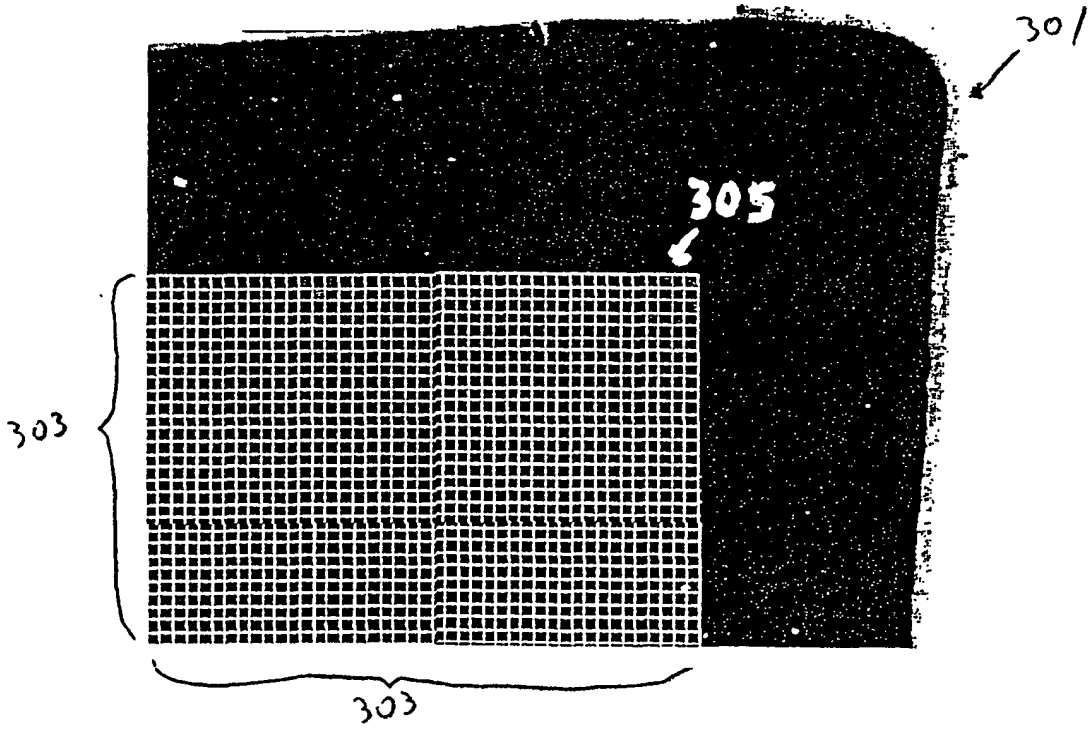


FIG. 7B

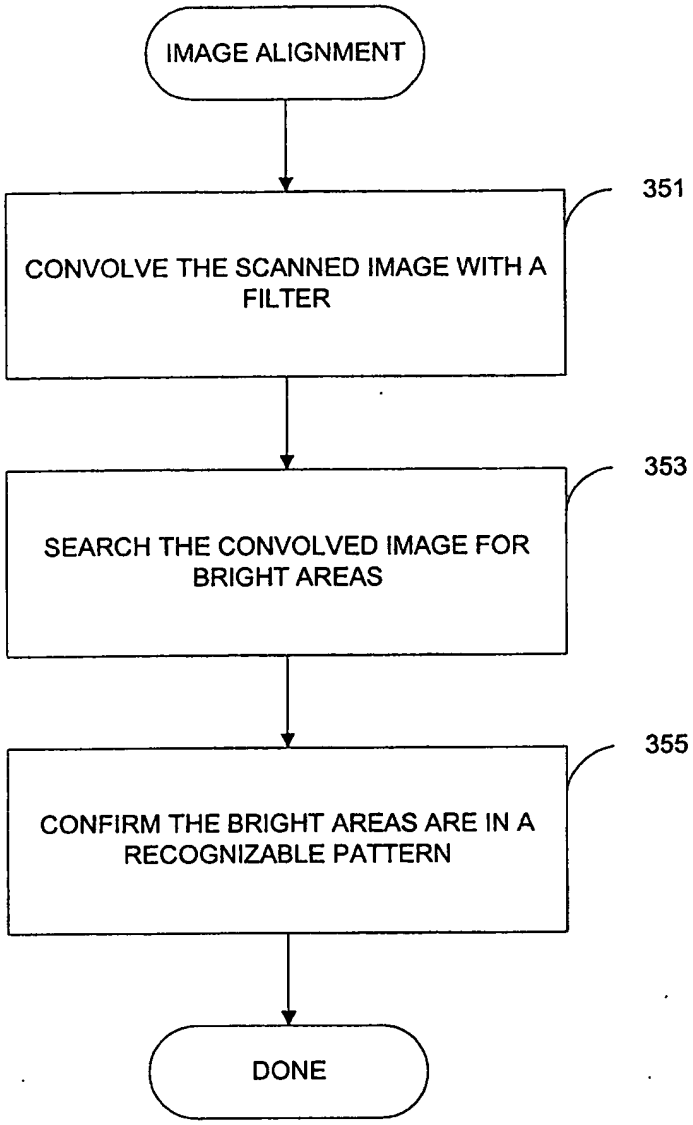


IMAGE ALIGNMENT
CONVOLVE THE SCANNED IMAGE WITH A FILTER
351
SEARCH THE CONVOLVED IMAGE FOR BRIGHT AREAS
353
CONFIRM THE BRIGHT AREAS ARE IN A RECOGNIZABLE PATTERN
355
DONE

FIG. 8

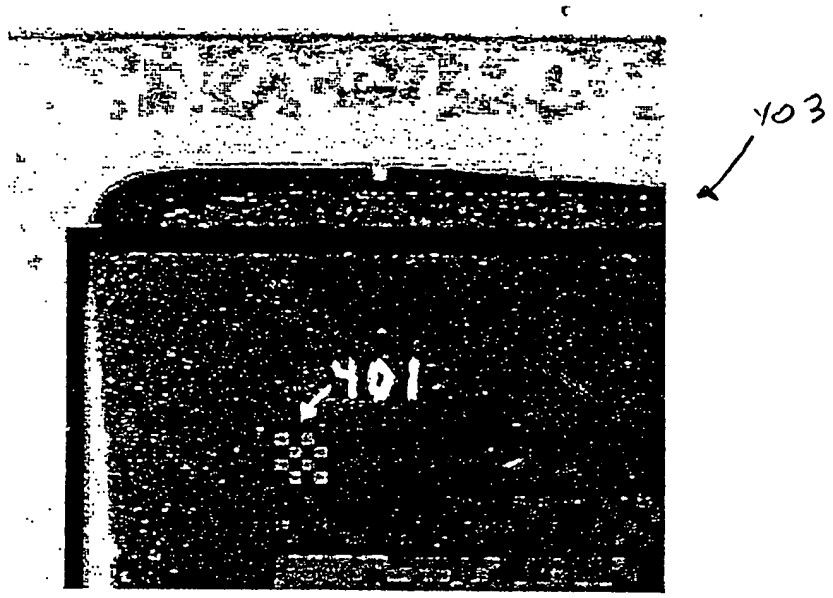


FIG. 9A

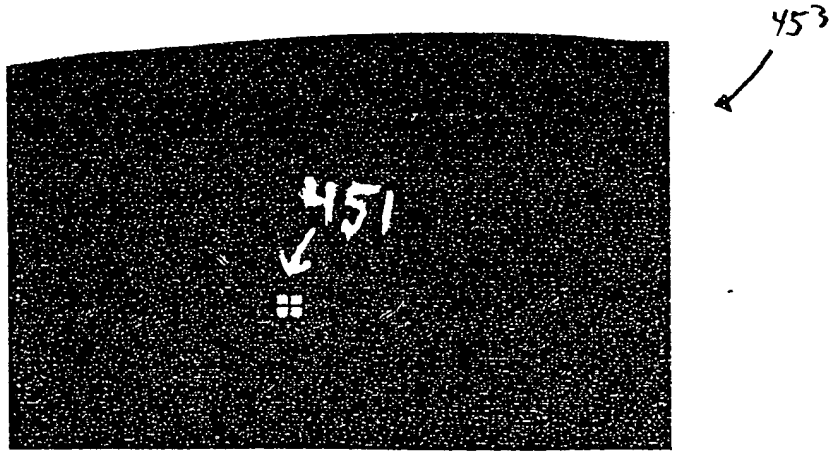


FIG. 9B

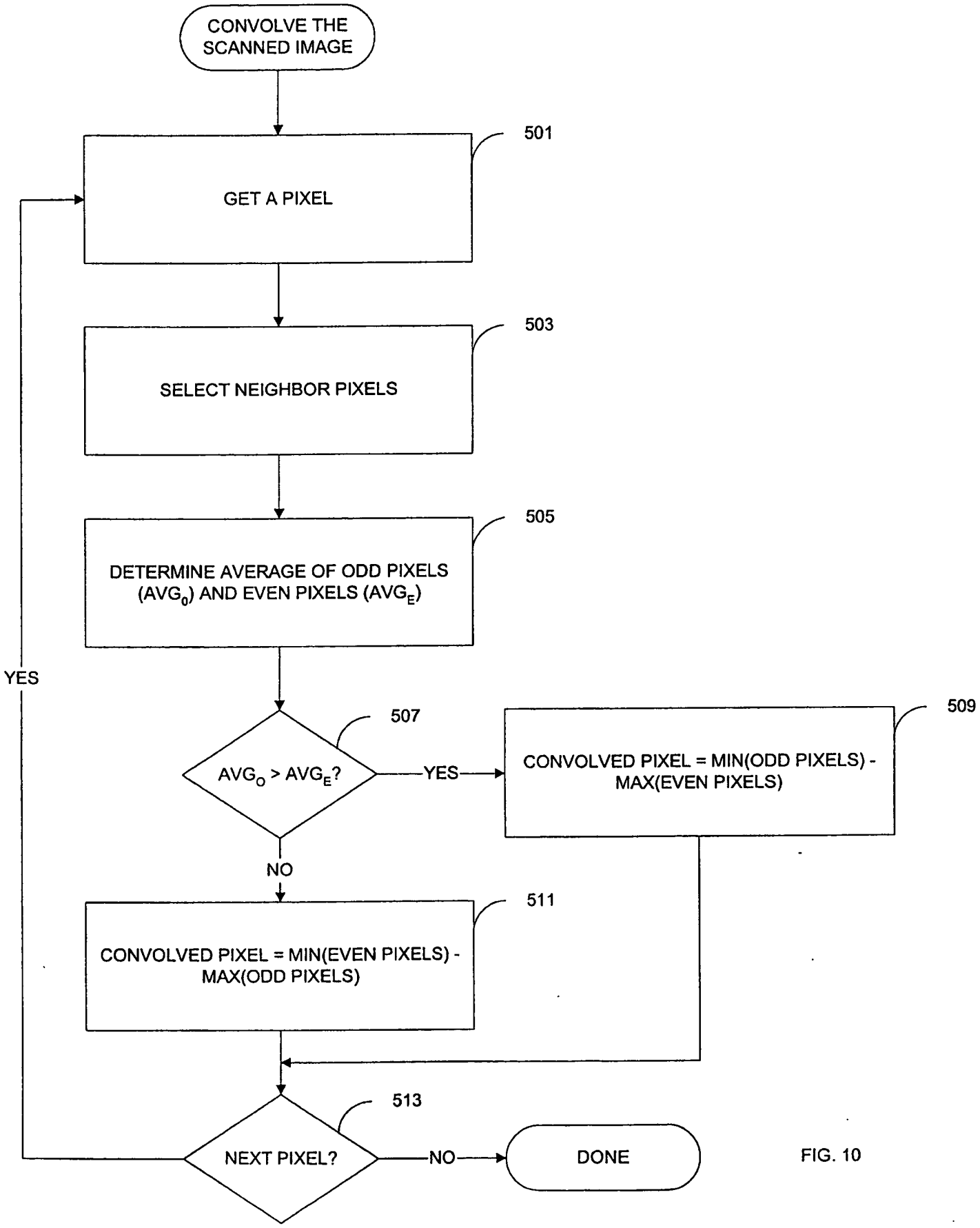

CONVOLVE THE SCANNED IMAGE
501
GET A PIXEL
503
SELECT NEIGHBOR PIXELS
505
DETERMINE AVERAGE OF ODD PIXELS (AVG_O) AND EVEN PIXELS (AVG_E)
507
$AVG_O > AVG_E$?
YES
509
CONVOLVED PIXEL = MIN(ODD PIXELS) - MAX(EVEN PIXELS)
NO
511
CONVOLVED PIXEL = MIN(EVEN PIXELS) - MAX(ODD PIXELS)
513
NEXT PIXEL?
YES
NO
DONE

FIG. 10

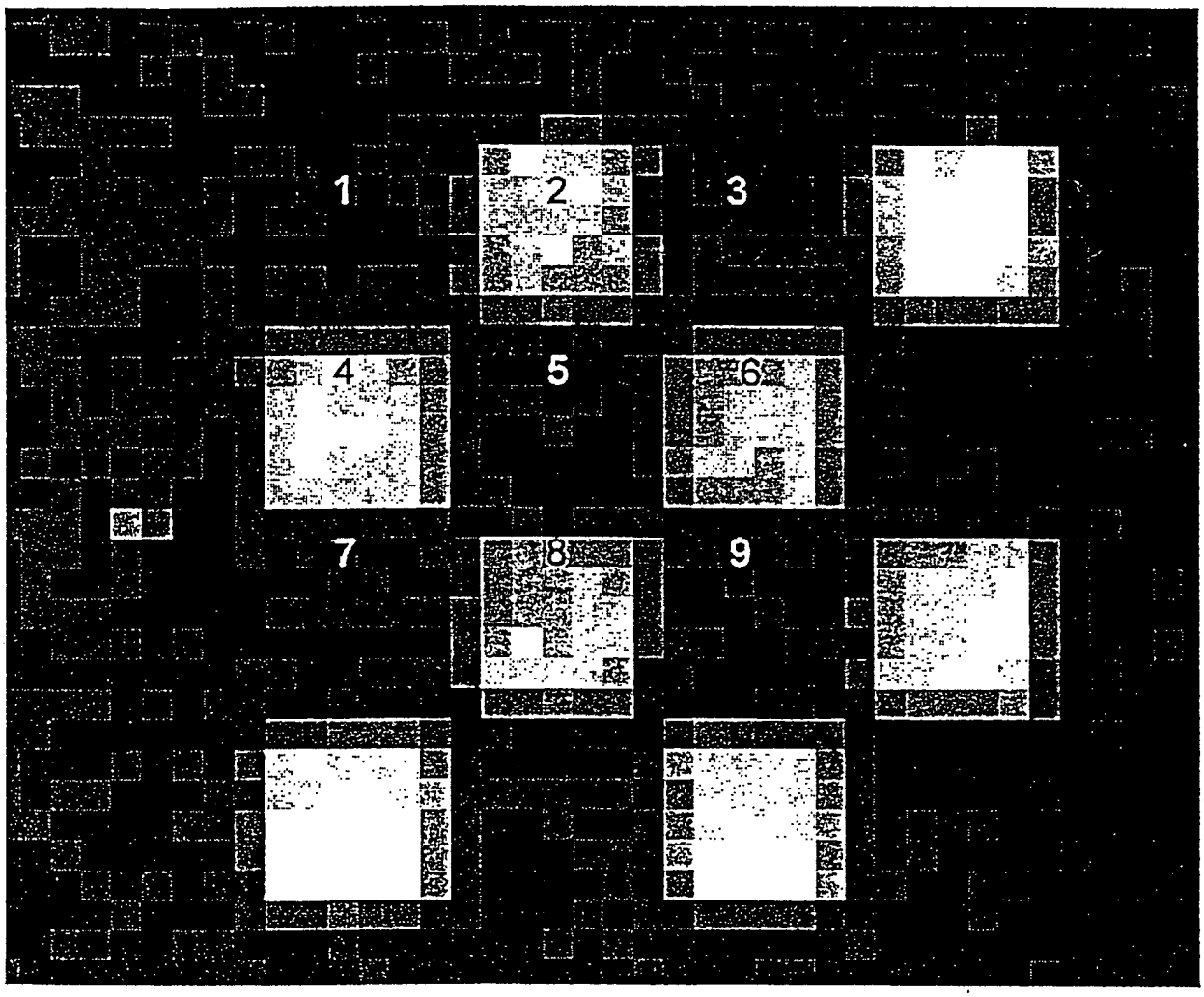


FIG. 11

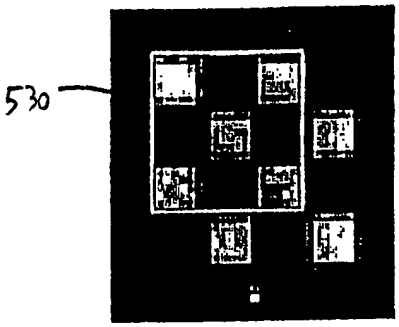


FIG. 12A

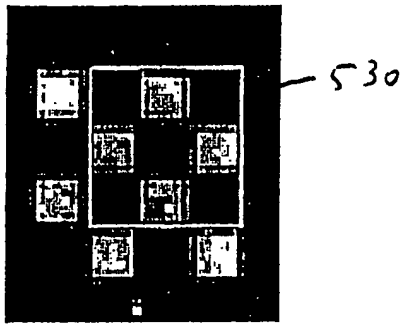


FIG. 12B

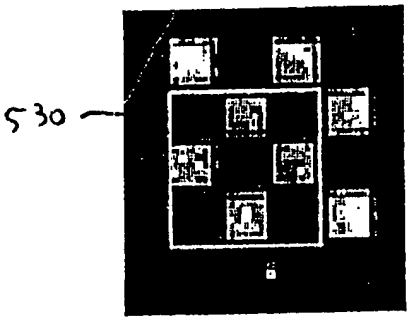


FIG. 12C

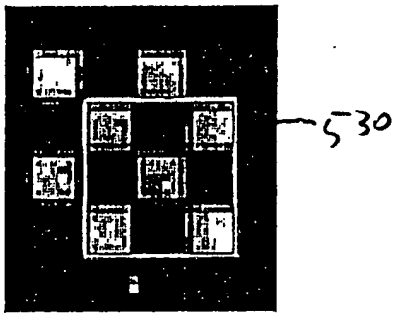


FIG. 12D

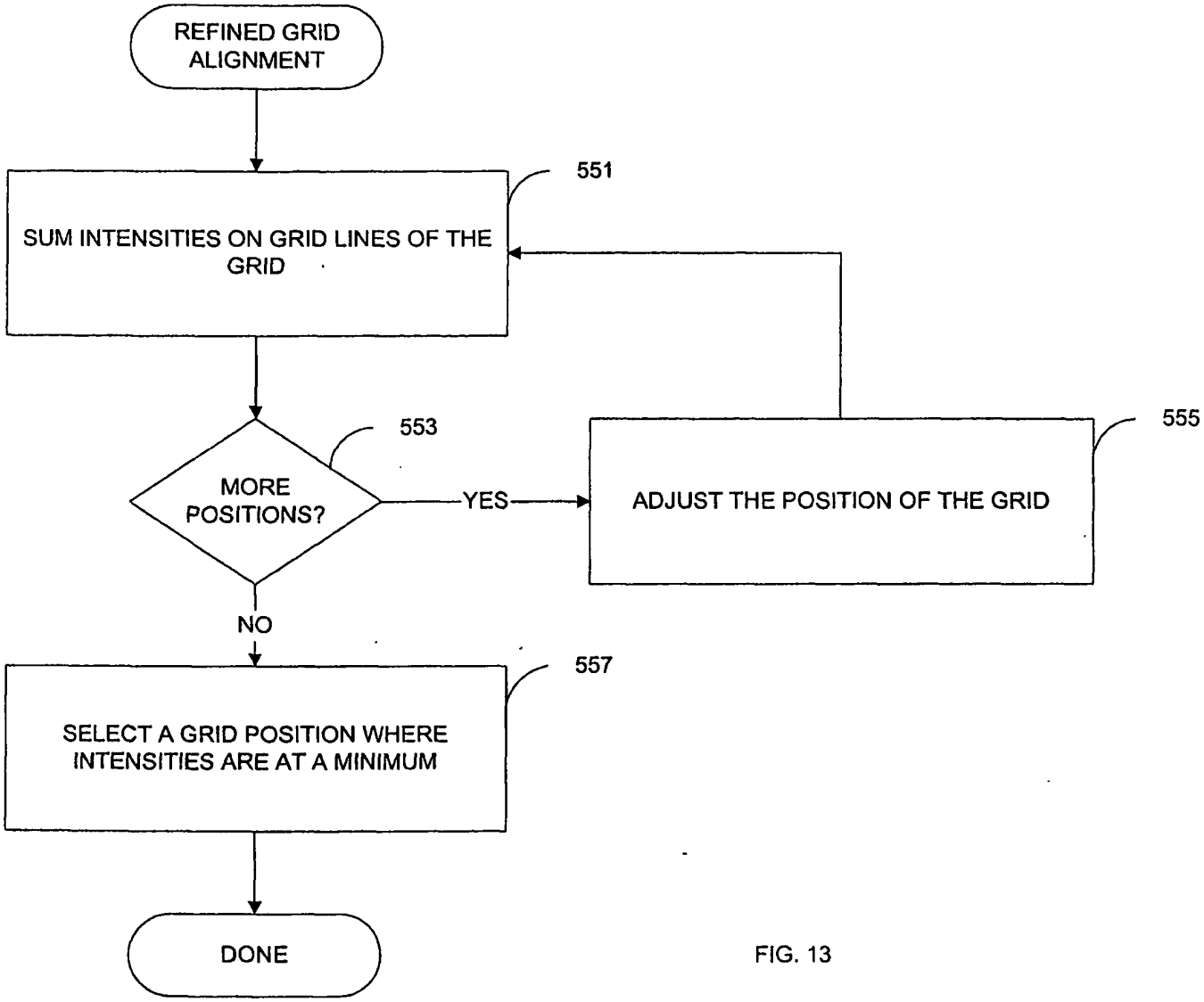

REFINED GRID ALIGNMENT
551
SUM INTENSITIES ON GRID LINES OF THE GRID
553
MORE POSITIONS?
YES
555
ADJUST THE POSITION OF THE GRID
NO
557
SELECT A GRID POSITION WHERE INTENSITIES ARE AT A MINIMUM
DONE

FIG. 13

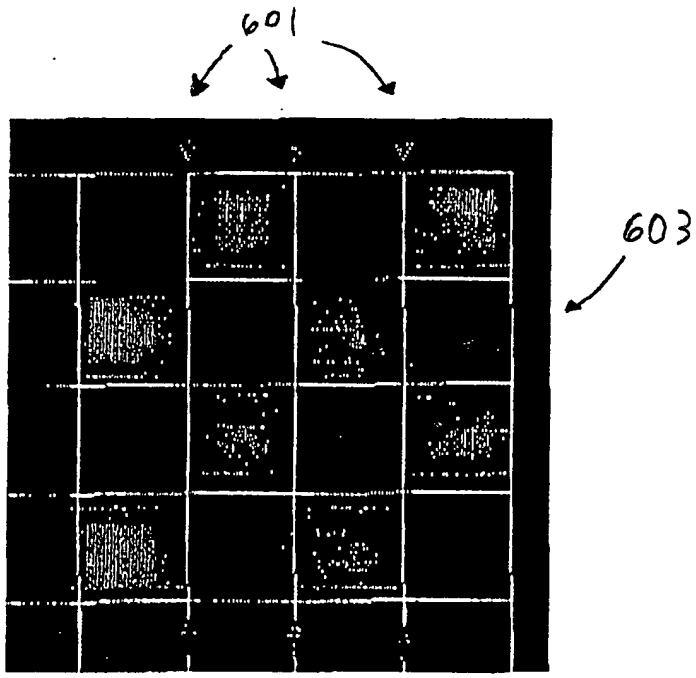


FIG. 14

PTO-103X
(Rev. 7-93)

FILING RECEIPT

UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office
ASSISTANT SECRETARY AND COMMISSIONER
OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

V/V

528 - 60

APPLICATION NUMBER	FILING DATE	GRP ART UNIT	FIL FEE REC'D	ATTORNEY DOCKET NO.	DRWGS	TOT CL	IND CL
08/195,889	02/10/94	1802	$840.00	11509117	12	10	1

VERN NORVIEL
TOWNSEND AND TOWNSEND KHOURIE AND CREW
STEUART STREET TOWER
ONE MARKET PLAZA, 20TH FLOOR
SAN FRANCISCO, CA 94105

Receipt is acknowledged of this patent application. It will be considered in its order and you will be notified as to the results of the examination. Be sure to provide the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION when inquiring about this application. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. If an error is noted on this Filing Receipt, please write to the Application Processing Division's Customer Correction Branch within 10 days of receipt. Please provide a copy of the Filing Receipt with the changes noted thereon.

Applicant(s) DAVID STERN, MT. VIEW, CA; PETER FIEKOWKSY, LOS ALTOS, CA.

FOREIGN FILING LICENSE GRANTED 07/26/94
TITLE
METHOD AND APPARATUS FOR DETECTION OF FLUORESCENTLY LABELED MATERIALS

PRELIMINARY CLASS: 436

(see reverse)

Attorney Docket No. 11509-117

PATENT APPLICATION

METHOD AND APPARATUS FOR DETECTION OF FLUORESCENTLY LABELED MATERIALS

Inventors: David Stern
Citizen of the United States
1285 Montecito Avenue
Apartment 36
Mountain View, CA 94043

Peter Fiekowski
Citizen of the United States
952 Springer Road
Los Altos, CA 94024

Assignee: Affymax Technologies, N.V.
Curacao, Netherland Antilles

TOWNSEND and TOWNSEND KHOURIE and CREW
Steuart Street Tower, 20th Floor
One Market
San Francisco, California 94105
(415) 543-9600

PATENT

Attorney Docket No. 11509-117

METHOD AND APPARATUS FOR DETECTION OF FLUORESCENTLY LABELED MATERIALS

BACKGROUND OF THE INVENTION

The invention provides a method and associated apparatus for detecting and analyzing reactions of fluorescently marked materials on a single substrate surface.

Certain macromolecules are known to interact and bind to other molecules having a very specific three-dimensional spatial and electronic distribution. Any large molecule having such specificity can be considered a target. The various molecules that targets selectively bind to are known as probes.

Methods and devices for detecting fluorescently marked targets on devices are known. Generally, the devices includes a microscope and a monochromatic or polychromatic light source for directing light at a substrate. A photon counter detects fluorescence from the substrate, while an x-y translation stage varies the location of the substrate. A computer controls the movement of the x-y translation table and data collection. Such devices are discussed in, for example, U.S. Pat No. 5,143,854 (Pirrung et al.) incorporated herein by reference for all purposes. See also PCT WO 92/10092 also incorporated herein by reference for all purposes.

Light from the light source is focused at the substrate surface by manually adjusting the microscope. Manual adjustment is, on occasion, time consuming and inconvenient. Moreover, due to inherent imperfections present in the x-y translation table and substrate, there is a possibility that the substrate will be out of focus as it is moved from one region to another. As a result, the data collected may be misrepresented.

Also, temperature sometimes impact a chemical reaction between targets and probes. Generally, targets are more active or form stronger bonds at lower temperatures while the converse is true at higher temperatures. However, if the temperature is too low, the binding affinity of the target may become excessively strong, thus causing target to bind with complements (matches) as well as non-compliments (mismatches). Hence, the ability to control temperature may affect optimum binding between the targets and probes while minimizing mismatches.

In addition, the microscope detection devices are uneconomical to use. Typically, these devices incorporates the use of a microscope, and a multichannel scaler, both of which are costly.

From the above, it is apparent that an improved method and apparatus for detecting fluorescently labeled targets on a substrate is desired.

SUMMARY OF THE INVENTION

Methods and devices for the detection of fluorescently labeled targets on a substrate are disclosed. The detection method and devices utilize a substrate having a large variety of probes at known locations on its surface. The substrate, when placed in a confocal detection device, is exposed to fluorescently labeled targets that bind to one or more of the probes.

The confocal detection device includes a monochromatic or polychromatic light source, means for directing an excitation light from the light source at the substrate, means for focusing the light on the substrate, means for controlling temperature of the substrate during a reaction, means for detecting fluorescence emitted by the targets in response to the excitation light by directing the fluorescence through confocal pinholes, and means for identifying the region where the fluorescence originated. The means for controlling the temperature may include a temperature controlled fluid filled flow cell. The means for

detecting the fluorescent emissions from the substrate, in some embodiments, include a photomultiplier tube. The means for focusing the excitation light to a point on the substrate and determining the region the fluorescence originated from may include an x-y-z translation table. Further, translation of the x-y-z table, temperature control and data collection are recorded and managed by an appropriately programmed digital computer.

In connection with one aspect of the invention, methods for analyzing the data collected by the fluorescent detection methods and devices are disclosed. Data analysis includes the steps of determining fluorescent intensity as a function of substrate position from the data collected; removing "outliers" (data deviating from a predetermined statistical distribution); and calculating the relative binding affinity of the targets from the remaining data. The resulting data are displayed as an image with the intensity in each region varying according to the binding affinity between targets and probes therein.

By using confocal optics, as well as focusing and temperature regulating techniques in conjunction with the data analysis methods, it is possible to quickly and accurately determine the relationship between structure and activity of certain molecules. Therefore, the potential for discovering novel probes with desirable pattern of specificity for biologically important targets is dramatically increased.

A further understanding of the nature and advantages of the inventions herein may be realized by reference to the remaining portions of the specification and the attached drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1a shows a detection system for locating fluorescent markers on the substrate;

FIG. 1b shows an alternative embodiment of a detection system for locating fluorescent markers on the substrate;

FIG. 1c shows another embodiment of a detection system for locating fluorescent markers on the substrate;

FIG. 2 is a flow chart illustrating the operation of the detection system;

FIG. 3 is another flow chart illustrating the focusing step of the detection system;

FIG. 4a is another flow chart illustrating the data acquisition step of the detection system;

FIG. 4b shows the relationship among the counters in the data acquisition board versus time.;

FIG. 5 is another flow chart illustrating the method of converting data representing photon counts as a function of position to data representing fluorescence intensity level as a function of position; and

FIG. 6 is another flow chart illustrating the data analysis step.

DESCRIPTION OF THE PREFERRED EMBODIMENT

CONTENTS

I. Definitions

The following terms are intended to have the following general meanings as they are used herein:

1. Complementary: Refers to the topological compatibility or matching together of interacting surfaces of a probe molecule and its target. Thus, the target and its probe can be described as complementary, and furthermore, the contact surface characteristics are complementary to each other.

2. Probe: A probe is a molecule that is recognized by a particular target. Examples of probes that can be investigated by this invention include, but are not restricted to, agonists and antagonists for cell membrane receptors, toxins and venoms, viral epitopes, hormones (e.g., opiates, steroids, etc.), hormone receptors, peptides, enzymes, enzyme substrates, cofactors, drugs, lectins, sugars, oligonucleotides, nucleic acids, oligosaccharides, proteins, and monoclonal antibodies.

3. Target: A target is a molecule that has an affinity for a given probe. Targets may be naturally-occurring or manmade molecules. Also, they can be employed in their unaltered state or as aggregates with other species. Targets may be attached, covalently or noncovalently, to a binding member, either directly or via a specific binding substance. Examples of targets which can be employed by this invention include, but are not restricted to, antibodies, cell membrane receptors, monoclonal antibodies and antisera reactive with specific antigenic determinants (such as on viruses, cells or other materials), drugs, polynucleotides, nucleic acids, peptides, cofactors, lectins, sugars, polysaccharides, cells, cellular

membranes, and organelles. Targets are sometimes referred to in the art as anti-probes. As the term targets is used herein, no difference in meaning is intended. A "Probe Target Pair" is formed when two macromolecules have combined through molecular recognition to form a complex.

II. Fluorescent Detection Device

Fig. 1a schematically illustrates a device used to detect fluorescently labeled targets on a substrate. Substrate 230 comprises a number of presynthesized probes on its surface 231. Such probes may be synthesized according to techniques described in U.S. Pat. No. 5,143,854 or PCT WO 92/10092 (attorney docket no. 11509-3900) or other techniques.

Substrate 230 is preferably transparent to a wide spectrum of light. In some embodiments, substrate 230 is made of a conventional microscope glass slide or cover slip. It is preferable that the substrate be as thin as possible while still providing adequate physical support. Preferably, the substrate is less than about 1 mm thick, more preferably less than 0.5 mm thick. Typically, the substrate is a microscope glass slide of about 0.7 mm or 700 μm thick. In alternative embodiments, the substrate may be made of quartz or silica.

Substrate 230 is mounted on a flow cell 220. Flow cell 220 is a body having a cavity 221 on a surface thereof. The cavity is between about 50 and 1500 μm deep with a preferred depth of 1000 μm. The bottom of the cavity is preferably light absorbing so as to prevent reflection of impinging light. In addition, the flow cell may be impervious to light.

When mounted to the flow cell, the substrate seals the cavity except for inlet port 223 and outlet port 224. According to a specific embodiment, the substrate is mounted to the flow cell by vacuum pressure generated from a vacuum pump 270. Optionally, one or more gaskets may be placed between the flow cell and substrate and the intervening space

is held at vacuum to ensure mating of the substrate to the gaskets.

Reagents, such as fluorescein labeled targets (fluorescence peak at about 530 nm) are injected into the cavity 221 through the inlet port 223 by a pump 240 or by using a syringe. The pump may be, for example, a Masterflex peristaltic pump made by Cole-Parmer Instrument Co or equivalent. Within the cavity, the reagents bind with one or more complementary probes on surface 231 of the substrate. The reagents are circulated into the cavity via inlet port 221 by and exit through the outlet port 224 for recirculation or disposal.

Flow cell 220 permits the substrate to remain in constant contact with reagents during detection, thereby allowing the substrate to be in equilibrium with targets therein. This arrangement also permits the user to manipulate test conditions without dismounting the substrate. In some embodiments, the flow cell provides means for controlling the temperature within the flow cell. The means for controlling temperature may be a recirculating bath device 260 that flows water through channels formed in the flow cell. In the specific embodiment, device 260 is a refrigerated circulating bath with a RS232 interface, catalog number 13270-615 distributed by VWR or equivalent. However, means such as a circulating air device, a resistance heater, a peltier device (thermoelectric cooler) or others may also be employed. Computer 190 monitors and controls device 260, thereby maintaining the flow cell at a desired temperature. Computer 190 may be selected from a wide variety of computers including, for example, a Gateway 486DX computer or a similar appropriately programmed computer.

Controlling the temperature in the flow cell is advantageous because temperature affects the chemical reaction between targets and probes. For example, the bond between the targets and probes is generally stronger at lower temperatures. However, if the temperature is too low, the binding affinity between targets and probes may become excessively strong so as to produce apparent (but erroneous)

matches. Thus, temperature can be controlled to maximize the binding affinity of complementary targets while minimizing mismatches.

Flow cell 220 is mounted on a x-y-z translation table 230. X represents the horizontal direction; y represents the vertical direction; and z represents the direction into and away from the microscope objective such that focusing may be performed. In some embodiments, the x-y-z translation table may be a Pacific Precision Laboratories Model ST-SL06R-B5M. Movement of the translation table is controlled by computer 190.

A light source 100 generates a beam of light to excite the fluorescein labeled targets in the flow cell. The light source may be a argon laser that generates a beam having a wavelength of about 488 nm, which in some embodiments may be a model 2017 or model 161C manufactured by Spectra-Physics. Other lasers, such as diode lasers, helium lasers, dye lasers, titanium sapphire lasers, Nd:YAG lasers or others may also be employed. The laser is directed at surface 231 through an optical train comprised of various optical elements which will be described below in detail.

The beam generated by laser 100 is typically nearly collimated and nearly Gaussian. However, a spatial filter may be optionally located in front of laser 100 to improve the Gaussian profile of the beam. The spatial filter may comprise of a lens 101, a confocal pinhole 103 and a lens 102. Lens 101 and 102, for example, may be 1/2" diameter 50 mm focal length anti-reflection coated plano convex glass lens or equivalent. Both lenses are configured such that both their back focal planes coincide with confocal pinhole 103. Pinhole 103, for example, may have a aperture of 30μm.

Thereafter, the light passes through a beam splitter 110 to a dichroic mirror 120. The beam splitter may be, for example, a non-polarizing 50% beam splitter cube made by Melles Griot model number 03BSC007 or equivalent while the dichroic mirror may be a LWP-45°S-488R/520T-1025 made by CVI Laser Corp. or equivalent. The functions of the beam splitter cube will later be described in more detail.

In some embodiments, dichroic mirror 120 passes light having a wavelength greater than about 520 nm, but reflects light having a wavelength of about 488 nm. Consequently, the 488 nm light from the laser is reflected by dichroic mirror 120 toward optical lens 130. Optical lens 130, in the specific embodiment, is 1/2" diameter -50mm focal length anti-reflection coated plano-concave glass lens made by Newport or equivalent. The light then passes through a microscope objective 140 to substrate 230 for magnification of the image sample. Microscope objective 140, in some embodiments, may be a 10x 0.3NA microscope objective, but other magnifications could also be used. In a preferred embodiment, the distance between lens 130 and microscope objective 140 is about 100 mm.

Microscope objective 140 focuses the light on surface 231, thereby exciting the fluorescein labeled targets. Preferably, the microscope objective produces a spot about 2 μm in diameter in its focal plane. The optical train described in the above embodiments produces a 2 μm diameter focal spot when used with a laser which generates a beam diameter of 1.4 mm, such as the Spectra-Physics model 2017.

In alternative embodiments, the 2 μm spot may be easily obtained when other types of light sources with different beam diameters are used. Since the diameter of the focal spot is inversely proportional to the diameter of the collimated beam produced by lens 102, the desired spot size may be achieved by varying the focal lengths of the spatial filter. Alternatively, a beam expander may be used to expand or compress the beam from the light source to obtain the desired spot size. For example, if a model 161C which generates a beam diameter of 0.7 mm, a 2 μm diameter focal spot may be achieved if the ratio of the lens in the spatial filter is 1:2 instead of 1:1. Thus, by varying the focal lengths of the lenses in the spatial filter and/or using a beam expander, the appropriate excitation spot size may be achieved from various beam diameters.

In a preferred embodiment, the 2 μm spot has a power of 50 μW. Depending on the light source used, a variable neutral density filter 310 may be inserted between the laser

100 and the optical train to attenuate the power of the laser to the desired power level.

In response to the excitation light, fluorescein labeled targets in the flow cell fluoresce light having a wavelength greater than about 520 nm. The fluorescence will be collected by the microscope objective 140 and passed to optical lens 130. Optical lens 130 collimates the fluorescence and passes it to dichroic mirror 120. In practice, light collected by microscope objective contains both fluorescence emitted by the fluorescein and 488 nm laser light reflected from the surface 231.

The laser component reflected from the substrate is reflected by dichroic mirror 120 back to beam splitter 110. Beam splitter 110 directs the laser component through a lens 175. The lens, in some embodiments, may be a 1/2" diameter 50 mm focal length anti-reflection coated plano convex glass lens made by Newport, but equivalent thereof may be used. Lens 175 focuses the laser component to a photodiode 170. Preferably, a confocal pinhole 171 is located between lens 175 and photodiode 170. Confocal pinhole transmits substantially only the reflected light originating from the focal plane of the microscope to photodiode 170 while reflected light originating from out-of-focus planes are blocked. In some embodiments confocal pinhole 171 has an aperture of 50 μm. Photodiode 170 generates a voltage corresponding to the intensity of the detected light. Photodiode may be, for example, a 13 DSI007 made by Melles Griot or equivalent, or other light detection devices, such as photomultiplier tube or avalanche photodiode may be used. Output from the detection device is used by computer 190 to focus the laser at a point on surface 231 of substrate 230.

As for the fluorescent component, most of it will pass through the dichroic mirror 120 since its wavelength is greater than about 520 nm. The fluoresced light is then focused by a lens 125 to a photomultiplier tube 160 for detecting the number of photons present therein. Lens 125, in a preferred embodiment, is a 1/2" diameter 50mm focal length anti-reflection coated plano convex glass lens made by

Newport, but equivalent lens may be used. A confocal pinhole 161 may be located adjacent to lens 125. Confocal pinhole transmits florescence originating from the focal plane and filters out light originating from other planes, such as from the glass or reagent. Accordingly, the signal-to-noise ratio of the fluoresced light is increased. Additionally, a filter 165 is preferably located between photomultiplier tube and confocal pinhole 161. In a specific embodiment, the filter transmits light having a wavelength greater than about 515 nm such as an Omega Optical 515 EFLP. In an alternative embodiment, the filter may transmit light having a wavelength between about 515 and 545 nm such as a 530 DF30 made by Omega Optical. Thus, photomultiplier tube 160 detects substantially only fluoresced light.

In the specific embodiment, photomultiplier tube 160 is a Hamamatsu R4457P photomultiplier tube with Hamamatsu C3866 preamplifier/discriminator. The Photomultiplier tube generates approximately a 2 mV pulse for each photon detected. Each of these 2 mV pulses are converted to a TTL pulse by the preamplifier/discriminator. The TTL pulses, each one corresponding to a photon detected by the photomultiplier tube, are then collected by a data acquisition board 210. The data acquisition board may be a National Instruments "Lab-PC+" or equivalent.

Data acquisition board 210, typically, contains an Intel 8254 or equivalent counter/timer chip. This chip contains three counters, counter 0, counter 1 and counter 2. Counter 0 controls the operations of counters 1 and 2 for collecting data. Preferably, counter 0 is programmed to generate a square wave with a period which is equal to twice the data acquisition time per pixel. The output of counter 0 is coupled to an external circuit board 200 which provides logic for inverting the square wave. In a preferred embodiment, the inverted output of counter 0 is connected to the gate input of counter 2 while the non-inverted output is connected to the gate input of counter 1.

In a preferred embodiment, the data acquisition board is not be able to read or store the fast 10 ns pulses

generated by preamplifier/discriminator (it is too fast for the 8254 chip). To solve this problem, external circuit board 200 may additionally provide means for slowing down the pulses. For example, the logic in external circuit board 200 may convert these pulses to 50 ns pulses with at least a 50 ns interval between pulses.

The output of the C3866 preamplifier/discriminator, via external circuit board 200, is connected to the clock inputs of counters 1 and 2. When counter 1 or counter 2 is gated on, it counts pulses generated by the preamplifier/discriminator; when it is gated off, it ceases to count and computer 190 reads the accumulated number of counts therein. After the computer reads the count from either counter 1 or 2, it is initialized to zero. Counter 1 or 2 is initialized on the first clock pulse after the gate input goes high. The initialization pulse is about a 50 ns pulse that is generated by the logic in the external circuit board 200 about 50 ns after each transition of the square wave signal from counter 0. The data stored in counter 1 or 2 represents the photon count as a function of substrate position.

After data are collected from a region of the substrate, substrate 230 is moved so that light can be directed at a different region on the substrate. The process is repeated until all regions on the substrate have been scanned. Generally, regions that contain a complementary probe will tend to exhibit a higher photon count than regions that do not contain a complementary probe.

Although the above embodiments have been described for use in detecting emissions of fluorescein excited by an 488 nm argon laser, it will be apparent to those skill in art that other dyes and excitation sources may used by simply modifying the elements in the optical train. For example, dichroic mirror 120 may be changed accordingly to pass light having a wavelength comparable to the fluorescence peak of the dye used, but reflect light from the excitation source. Also, filter 165 is changed to pass substantially only light having a wavelength similar to the fluorescence peak of the dye used.

In this manner, the detection device can be easily modified to accommodate other types of excitation light and/or dyes.

Fig. 1b illustrates an alternative embodiment of the fluorescence detection device shown in Fig. 1a. Fig. 1b is similar to the one shown in Fig. 1a and the common elements have been numbered with the same reference numerals. The main difference between this embodiment and that of Fig. 1a is that a photodiode 180 is provided to detect a component of the light generated by laser 100. Light generated by the laser, as in Fig. 1a, is directed at the beam splitter. However, a component of this light is directed to photodiode 180. Photodiode 180 generates a voltage value representing the intensity of the light. This voltage signal is used by the computer 190 to monitor and control the intensity of the laser.

Fig. 1c illustrates an alternative embodiment of the fluorescence detection device. Fig. 1c is similar to the embodiment shown in Fig. 1a and the common elements have been numbered with the same reference numerals. However, the embodiment in Fig. 1c provides means for detecting a second fluorescent color. Two color detection is required when two different types of targets, each labeled with a different dye, are exposed to a substrate synthesized with probes. In some embodiments, fluorescein and rhodamine dyes may be used to label two different types of targets respectively. Typically, each dye will have a fluorescence peak at different wavelengths. For example, the fluorescence peak of fluorescein is about 530 nm while that of a typical rhodamine dye is about 580 nm.

To detect the second fluorescent color, a second dichroic mirror 300 is employed. If rhodamine and fluorescein were used, then dichroic mirror 300 is designed to pass light having a wavelength greater than about 570 nm (rhodamine emissions) and reflect light having a wavelength less than about 560 nm (fluorescein emissions). Light with a wavelength less than 560 nm is reflected to a lens 126 and through a confocal pinhole 151. Lens 126, may be equivalent to lens 125 while confocal pinhole 151 may be similar to confocal pinhole

161. Filter 155 then filters the less than 560 nm light before entering a second photomultiplier tube 150. Filter 155 may be an Omega Optical 530DF30 or equivalent that passes light with a wavelength between about 515-545 nm. This ensures that substantially only fluorescein emissions are detected by the photomultiplier 150.

On the other hand, light having a wavelength greater than 570 nm passes through dichroic mirror 300 to a lens 125. Lens 125 then directs the greater than 570 nm light through a pinhole 161 and a filter 165 before entering photomultiplier tube 160. Filter 165 may be a Schott OG570 or equivalent which passes light having a wavelength greater than 570 nm, thereby ensuring substantially only rhodamine emissions are detected by photomultiplier 160.

Output of the preamplifier/discriminator from the photomultiplier tube 150 is processed by the external circuit board 200 before being connected to counters 1b and 2b on the data acquisition board 205. Data collection by counters 1b and 2b are controlled by counter 0 from data acquisition board 210 which generates a square wave to the gated inputs of the counters 1 and 2 via the external circuit board 200, similar to that of counters 1 and 2 on data acquisition board 210.

According to the embodiment in Fig. 1c, two fluorescence color can be detected by employing a second dichroic mirror, photomultiplier tube and associated lens, confocal pinhole and filter. The embodiment illustrated in Fig. 1c may be expanded by one skill in the art to detect more than two fluorescence colors by employing an additional dichroic mirror, photomultiplier tube and associated lens, confocal pinhole and filter for each additional fluorescence color to be detected.

III. Details on the Operation of a Fluorescent Detection Device

In the specific embodiment, data are acquired continuously along a line which is broken up into data points or pixels. The pixel size preferably ranges from 1 to 100 μm.

Since pixels are preferably square, the distance between scan lines will generally be between 1 to 100 μm. Each of these pixels will be used to form a digital image. The image resolution (i.e., the degree of discernable detail) depends upon the number of pixels that forms the image. The greater the number of pixels, the closer the digitized data will approximate the original image. Therefore, it is desirable to have as many pixels as possible. However, the number of pixels is directly related to the length of time required to scan the sample. In practice, it is found that having about 16 to 100 pixels per feature (or about 4-10 pixels along the length of the feature) is preferable to achieve the desire resolution in a reasonable amount of time.

The number of photons that is detected in each pixel is contingent upon several factors. The most obvious factor, of course, is the amount of fluorescently labeled targets present in the pixel. Other, but not so obvious, factors include the intensity of the excitation light, length of time that the targets are excited, and quantum efficiency of the photocathode at the fluorescence emission wavelength.

For example, exciting a region of about 2 μm with 50 μW of power will yield approximately 1600 W/cm^2 or 3.9×10^{21} photons/(sec cm^2). At this intensity, the fluorescence rate is $Q_e k_a k_f/(k_a+k_f) = 1.1 \times 10^6$/sec (see Table I) with a photodestruction rate of $Q_b k_a k_f/(k_a+k_f) = 36$/sec. In a typical substrate synthesized with polymer sequences at about 5 nm apart, approximately 4×10^4 molecules/μm^2 or 1.25×10^5 molecules will be present in the excitation volume. If it is estimated that about 1% of these sequences bind with fluorescein labeled targets, then about 1250 molecules are excited or 1.4×10^9 fluorescence photons are generated per second. However, in a typical detection device, only about 2% of these photons are collected by the microscope objective while about 98% never even make it into the optical train of the detection device. Of the 2% collected, about 50% are lost in the optical train and of the remaining photons, only about 10% are detected by the photomultiplier tube due to quantum efficiency of the photocathode at the fluorescein emission wavelength. Thus,

approximately 1.4×10^6 photons might be counted per second. From the above, it is apparent that these factors affect the number of photons detected at each pixel.

TABLE I

Fluorescein Optical Parameters

[Tsien, R.Y., Waggoner, A. Fluorophores for confocal microscopy: photophysics and photochemistry. In *Handbook of Biological Confocal Microscopy*; Pawley, J., Ed.; Plenum Press: New York, 1990; pp. 169-178]

Absorption cross section, σ, $cm^2 molecules^{-1}$	3.06×10^{-16}
Fluorescence rate constant, k_f, s^{-1}	2.2×10^8
Absorption rate constant (1600 W/cm^2) k_a, s^{-1}	1.2×10^6
Fluorescence quantum efficiency, Q_e	0.9
Photodestruction efficiency, Q_b	3×10^{-5}

In the present invention, it is preferable to detect as many photons as possible, preferably about 1000 photons per pixel in pixels containing complementary probes and targets because the signal-to-noise ratio is equal to the square root of the number of photon counts. Since the number of photons collected will vary according to the power and length of time the targets are excited by the light, the signal-to-noise ratio may be improved by increasing the laser power or length of time the substrate is exposed to the laser or a combination thereof.

Figs. 2-6 are flow charts describing a specific embodiment. Fig. 2 is an overall description of the system's operation. Referring to Fig 2, the detection system is initialized at step 200. At step 201, the system prompts the user for test parameters such as:

a) temperature of the substrate;

b) number of scans to be performed;

c) time between scans;

d) refocus between scans;

e) pixel size;

f) scan area; and

g) scan speed.

The temperature parameter controls the temperature at which detection is performed. Temperature may vary depending on the type of polymers being tested. Preferably, testing is done at a temperature that produces maximum binding affinity while minimizing mismatches.

The number of scan parameter corresponds to the number of times the user wishes to scan the substrate while the time between scans parameter controls the amount of time to wait before commencing a subsequent scan. In this manner, the user may perform a series of scan and if desired, each at a different temperature. Preferably, the time between scans is chosen to allow the system to reach chemical equilibrium before commencing a subsequent scan.

In an alternative embodiment, means may be provided to increase the temperature at set increments automatically after each scan. Further, the user may optionally choose to refocus the substrate after each scan.

The pixel size parameter dictates the size of each data collection point. Generally, the size is chosen which results in at least 16 data collection points or pixels per synthesis region ("feature").

Scan area parameter corresponds to the size of the substrate to be tested. Scan speed parameter sets the length of time the laser excites each pixel. The slower the speed, the higher the excitation energy per pixel which will result in higher fluorescence count per pixel. Thus, increasing the laser power or decreasing the scan speed or a combination thereof will increase the photon count in each pixel. Typically, the scan rate is set to generate approximately a count of 1000 photons for pixels having fluorescently-marked targets.

At step 202, the system initializes the x-y-z translation table by locating the x-y-z stages at their home position. At step 203, the system focuses the laser on the

surface 231 of the substrate. At step 204, the system locates the x-y-z table at its start position. At step 205, the system begins to translate the vertical stage, thereby collecting a series of data points over a vertical line at step 206. When a line of pixels has been scanned at step 207, the x-y-z translation table moves the horizontal stage to collect data from the next line of pixels at step 208. The collected data is written to the file as the substrate is repositioned at the top of the next line. Steps 205 through 208 are repeated until data from all regions have been collected. At step 209, the system determines if there are any more scans to be performed according to the set up parameters. If there are, the system at steps 210 and 211 determines the amount of time to wait before commencing the next scan and to either repeat the process from step 203 (if refocusing of the substrate is desired) or 204. Otherwise, the scan is terminated.

Fig. 3 illustrates the focusing step 203 in greater detail. Auto-focusing is accomplished by the system in three phases. In the first phase, the system focuses the laser roughly on the back surface of the substrate. At step 301, the laser is directed at one corner of the substrate. After the light contacts the surface, it is reflected back through microscope objective 140 and optical lens 130 to dichroic mirror 120. The dichroic mirror reflects the light to beam splitter 110 which guides the beam to lens 175. Lens 175 focuses the light through a pinhole 171 to photodiode 170. Pinhole 171 is a confocal pinhole, which transmits light that is reflected from the surface located at the focal plane and blocks light that is reflected or scattered from other surfaces. At step 302, photo-diode 170 generates a voltage corresponding to the intensity of the reflected light. The flow cell is then moved about 10 microns closer to the microscope objective at step 315, and the process from step 302 is repeated. The loop commencing at step 302 is repeated until the voltage generated by the photodiode has peaked (i.e., the present voltage value is less than the previous voltage value), at which time, the laser is roughly focused on

the backside of the substrate. Since 10 μm steps are taken, the light is focused slightly inside the substrate.

At step 304, the system continues with the next focusing phase by moving the flow cell closer to the microscope objective. In a preferred embodiment, the distance over which the flow cell is moved is about equal to half the thickness of the substrate. The default distance is 350 mm (1/2 the thickness of a typical substrate used) or a distance entered by the user representing half the thickness of the substrate used. Again, at step 305, photo-diode 170 generates a voltage corresponding to the intensity of the reflected light. Preferably, the flow cell is then moved about 10 microns closer at step 316, and the process from step 305 is repeated. The loop commencing at step 305 is repeated until the voltage generated by the photodiode has peaked, at which time, the laser is roughly focused on a point beyond surface 231.

At step 307, the flow cell is moved farther from the microscope objective in steps of about 1 μm. The computer reads and stores the voltage generated by the photodiode at step 308. At step 309, the encoder indicating the position of the focus stage is read and the resulting data is stored. The encoder determines the location of the focus stage to within about 1 micron. At step 310, the system determines if the present voltage value is greater then the previous value. If it is, the flow cell is then moved about 1 micron farther at step 317. Due to the presence of noise, the process from step 308 is repeated even after the voltage value has peaked. Specifically, the process is continues until the voltage generated by the photodiode 104 is less than the peak voltage minus twice the typical peak-to-peak photodiode voltage noise. At step 311, the data points are fitted into a parabola where x = encoder position and y = voltage corresponding to the position. Fitting the data points to a parabola gives a more accurate method of focusing than by merely taking the highest point due to the presence of noise in the data. At step 312, the system determines the focus position on surface 231 which corresponds to the maximum of the parabola.

At step 313, the system determines whether all four corners have been focused. If they have been, then the process proceeds to step 314. Otherwise, the system, at step 318 moves the x-y-z translation stage for directing the light at the next corner and returns to step 301. The focussing process beginning at step 301 is repeated until the all four corners of the substrate has been focused.

At step 314, the system assumes that the substrate is planar. Thus, the focus position of other pixels are obtained by interpolating the values collected from the focusing process. In this manner, auto-focusing for each pixel of the substrate is achieved.

By using the focusing method disclosed herein, the laser is focused on surface 231 of the substrate, which is significantly less reflective than the backside of the substrate. Generally, it is difficult to focus on a weakly reflective surface in the vicinity of a strongly reflective surface. However, this problem is solved by the present invention and thus, more accurate reading of the fluorescently marked targets is achieved.

Fig. 4a illustrates the data acquisition process beginning at step 205 in greater detail. In a specific embodiment, data are collected by repeatedly scanning the substrate in vertical lines until the sample is completely scanned. However, other techniques such as repeatedly scanning the substrate in horizontal lines, bidirectional scanning (acquiring data in both directions) or others may be employed.

At step 401, the x-y-z translation table is initialized at the start position. At step 402, the system calculates the constant speed at which the vertical stage is to be moved. The speed of the vertical stage is determined from the scan speed information entered by the user. Typically the scan speed is about 10 to 30 mm/sec or a speed at which a photon count of about 1000 photons will be generated for pixels having complementary probes. At step 403, the system calculates the constant speed at which the focusing stage is to be moved in order to maintain the

substrate surface 231 in focus. This speed is derived from the data obtained during the focusing phase.

At step 404, the system calculates the number of pixels per line by dividing the length of the scan line by the pixel size. At step 405, the system initializes counter 0 on the data acquisition board with a value such that a square wave having a period that is equal to twice the data acquisition time per pixel is generated. The period is calculated by dividing the pixel size information entered by the user at step 209 by the speed of the vertical stage derived from step 402. Counter 0 counts down until it reaches zero, at which time, a square wave transition has occurred, i.e., value of the square wave goes from low to high or high to low. Simultaneously, counter 0 is re-initialized with its initial value.

In a preferred embodiment, counter 1 on the data acquisition board is configured to store the photon counts from the even pixels while counter 2 is configured to store counts from the odd pixels, but other configurations may be employed. At step 406, counter 2 is initialized to zero by the rising edge of the first period of the square wave. Thereafter, counter 2 is enabled and begins to collect data at step 407.

The system, at step 408, polls counter 0 and compares the present value of counter 0 with its previous value. If the present value is less than the previous value, then counter 2 continues to accumulate photon counts. On the other hand, if the present value in counter 0 is greater than its previous value, a square wave transition has occurred. The falling edge of the square wave disables counter 2 from counting, thus, completing the scan of the first pixel. Simultaneously, counter 1 is initialized because its gate input is coupled to the inverted output of counter 0. The operation of counter 1 will be described in more detail during the discussion on the second pass of the loop beginning at step 406.

While counter 2 is disabled, the photon count stored in counter 2 is read at step 409. At step 410, the data is

written and stored in memory, for example, in the form of a data structure, an array, a table or other listing means. In the alternative, the data may also be written to a data file. At step 411, the system determines if there are more pixels in the line to scan. If there are, process repeats the steps beginning at 406. Otherwise, the system proceeds to step 412.

On the second pass of the loop beginning at step 405, the inverted falling edge (rising edge) of the square wave initializes and enables counter 1 to collect data at steps 406 and 407 respectively. At step 408, the inverted rising edge (falling edge) of the square wave disables counter 1 and data therein is read at step 409 and written to the computer at step 410.

In the specific embodiment, the counters are preferably configured to collect and store data alternately, i.e., when counter 1 collects data, counter 2 stores data. Fig. 4b illustrates the relationship among the count value in counter 0, the square wave, counter 1 and counter 2 versus time.

At step 411, the system determines if there are more pixels left to scan. The loop from step 405 through step 411 is repeated until all pixels in the line have been scanned. After each line has been scanned, the system at step 412 calculates a gray scale for imaging. In a preferred embodiment, the gray scale contains 16 gray levels, but a scale of 64, 256 or other gray levels may be used. In alternative embodiments, a color scale may be used instead of a gray scale. Preferably, the middle of the scale corresponds to the average count value detected during the scan. Thereafter, the raw data points are stored in a data file at step 413. At step 414, the scaled data representing an image of the scanned substrate regions may be displayed on a screen or video display mean in varying shades of gray or colors. Each shade or color corresponds to the intensity level of fluorescence at the respective regions.

While the image of the previous scanned line is being displayed, the system determines if there are any more lines to scan at step 415. If so, the horizontal stage is

translated in preparation for scanning the next line at step 416. The distance over which the horizontal stage is moved is equal to about 1 pixel. Simultaneously, the vertical stage is moved to the top of the next scan line. Thereafter, the system repeats the process starting at step 405 for the next scan line. The loop from step 405 to step 415 is repeated until scanning of the substrate is completed. In this manner, the system simultaneously displays and collects data. Upon completion, the system creates a data file wherein the data represents an array of photon counts as a function of substrate position.

By counting the number of photons generated in a given area in response to the excitation light, it is possible to determine where fluorescently marked molecules are located on the substrate. Consequently, it is possible to determine which of the probes within a matrix of probes is complementary to a fluorescently marked target.

According to preferred embodiments, the intensity and duration of the light applied to the substrate is controlled by the computer according to the set up parameters entered at step 201. By varying the laser power and scan stage rate, the signal-to-noise ratio may be improved by maximizing fluorescence emissions. As a result, the present invention can detect the presence or absence of a target on a probe as well as determine the relative binding affinity of targets to a variety of sequences.

In practice it is found that a target will bind to several peptide sequences in an array, but will bind much more strongly to some sequences than others. Strong binding affinity will be evidenced herein by a strong fluorescence signal since many target molecules will bind to that probe. Conversely, a weak binding affinity will be evidenced by a weak fluorescence signal due to the relatively small number of target molecules which bind in a particular probe. As such, it becomes possible to determine relative binding avidity (or affinity in the case of univalent interactions) of a probe herein as indicated by the intensity of a fluorescent signal in a region containing that probe.

Semiquantitative data on affinities may also be obtained by varying set up conditions and concentration of the targets in the reagent. This may be done by comparing the results to those of known probe/target pairs.

While the detection apparatus has been illustrated primarily herein with regard to the detection of marked targets, the invention will find application in other areas. For example, the detection apparatus disclosed herein could be used in the fields of catalysis, DNA or protein gel scanning, and the like.

IV. Data Analysis System to Determine Relative Binding Strength of Targets

Before the data file representing an array of photon counts as a function of position is analyzed to determine the relative binding affinity of targets, the data file is preferably converted to an image file wherein the data is indicative of fluorescence intensity level as a function of substrate position.

Fig. 5 illustrates the process for converting or scaling the data from photon counts to fluorescence intensity level in greater detail. The conversion procedure is started by prompting the operator for the name of data file of interest. At step 501, the system retrieves the specified data file for analysis.

At step 502, the user directs the system to scale the data either in relative range or absolute range mode. In the absolute range mode, the data is scaled by using the set of values between the minimum and maximum number of photons. If absolute range mode is chosen, the system proceeds to step 504 which will later be described.

On the other hand, if relative range mode is chosen, the system proceeds to step 503. Scaling the data in relative range is advantageous. Particularly, relative range mode minimizes the effect of aberrations resulting from dirt or other imperfections on the substrate by ignoring a certain percentage of data points at the intensity extremes. From

experience, this range is typically from about 1% of the lowest values and 0.03% of the highest values. At step 503, the user enters the range values or in the alternative, the system may provide for default values. For example, if the user enters .03% and 1% as the relative range values and there are 100,000 pixels in the image, the brightest 30 pixels (0.003x100,000) and dimmest 1000 (0.1x100,000) pixels are clipped (clamped to its next highest or lowest value respectively).

At step 504, the system determines if the dynamic range (i.e., the ratio of the count in the brightest pixel and the dimmest pixel) of the image is high. If the ratio is high, the system prepares the data for logarithmic scaling at step 505. Otherwise, the system proceeds to step 506 where the data will be scaled linearly. In alternative embodiments, the system may prepare the data for square root scaling instead of logarithmic scaling. In this manner, lost of valid data at the low intensities having low photon counts are avoided, thereby, increasing the resolution at these intensities.

At step 506, the system obtains the minimum and maximum photon counts from the data. The range between these two count values is used to scale the data into digitized units. Preferably, the digitized units have a range from 0 to 255, each one representing a gray level or a color. Alternatively, other ranges such as from 0 to 63 may be used. Generally, the average photon count value is placed in the middle of scale.

Upon completion of the conversion process, an image file representing fluorescence intensity is created and stored in memory at step 507. At step 508, the system may optionally display the image file. The intensity level of the displayed image varies from region to region according to the binding affinity of the targets to the polymer sequence therein. The brightest signals typically represent the greatest binding affinity while signals of lesser intensity represent lesser degrees of binding affinity.

As described, data are collected over regions substantially smaller than the area in which a given polymer or feature is synthesized. For example, the length of a pixel is generally 1/4 to 1/10 the length of a feature (or the area of a pixel is 1/16 to 1/100 the area of a feature). Hence, within any given feature, a large number of fluorescence data points or pixels are collected.

A plot of the number of pixels versus the fluorescence intensity for a scan of a substrate synthesized with probes when it has been exposed to, for example, a labeled antibody will typically take the form of a bell curve. However, spurious data are observed, particularly at higher intensities. Since it is preferable to used an average of the fluorescence intensity over a given synthesis region in determining the relative binding affinity, these spurious data points will tend to undesirably skew the data.

Fig. 6 illustrates one embodiment of the of a system which provides for the removal of these undesirable spurious data points as well as the determination of the relative binding efficiency of the sample from an average of the remaining data points.

Referring to Fig. 6, the system is initialized by requesting the user to enter the name of a image file of interest. At step 601, the system retrieves the image file and prompts the user to enter the four corners of the image at step 602. Next, at steps 603 and 604, the system prompts the user for the number of cells located horizontally and vertically on the substrate. From the information entered by the user and the image file, the system creates a computer representation of a histogram for each cell at step 605. The histogram (at least in the form of a computer file) plots the number of pixels versus intensity.

At step 606, the main data analysis loop is performed for each synthesis site. Analyzing the histogram for the respective synthesis site, the system calculates the total intensity level and number of pixels for the bandwidth centered around varying intensity levels. For example, as shown in the plots to the right of step 606, the system

calculates the number of pixels in the bandwidth using boxcar averaging technique. This process is then repeated until the entire range of intensities have been scanned. At step 607, the system determines which band has the highest total number of pixels. The data from this band is used to derive statistical data for each synthesis site. The statistical data include the peak value, mean intensity and standard deviation of intensity level. Thus, data that are beyond this band are excluded from the statistical analysis. Assuming the bandwidth is selected to be reasonably small, this procedure will have the effect of eliminating spurious data located at both the higher and lower intensity levels. This loop is repeated until all the cells have been processed.

At step 610, an image in the form of a grid representing the substrate is displayed. Each block in the grid represents a region synthesized with a polymer sequence. The image intensity of each region will vary according to the binding affinity between the polymer sequence and targets therein. Statistical data, such as the peak and average intensity corresponding to each region are also displayed.

To improve imaging, pixels located at transitions between synthesis regions are ignored. The image, in some instances, requires only one pixel space between the cells when the transition of the intensity between the synthesis regions is sharp and distinct. However, if the transition is fuzzy or smeared, the user, at step 611, can select a greater number of pixel spaces between the cells to increase image resolution. If the user enters a value indicating a greater number of pixel spaces is desired, the system at step 616 reformats the image accordingly.

At step 612, the system retrieves the file created during the synthesis process of the substrate being analyzed. The synthesis file contains sequence information as a function of location. The system integrates the synthesis file with the image file and sorts the data therein. Through this process, the molecular sequence of complementary probes and the intensity as a function of location is available.

Further, the user, at step 614, may analyze a specific synthesis region within the grid. If instructed, the system will display the corresponding substrate position, number of photons, number of pixels and the molecular sequence at that synthesis site. The data analysis software also provides the user with many functions which are common to image processing, such as magnification and image enhancement.

V. Conclusion:

The present invention provides greatly improved methods and apparatus for detection of intensity of fluorescence on a substrate. It is to be understood that the above description is intended to be illustrative and not restrictive. Many embodiments will be apparent to those skill in the art upon reviewing the above description. Although the detection apparatus has been illustrated primarily herein with regard to the detection of marked targets, it will readily find application in other areas. For example, the detection apparatus disclosed herein could be used in the fields of catalysis, DNA or protein gel scanning, and the like. The scope of the invention should, therefore, be determined not with the reference to the above description, but should instead be determined with reference to the appended claims, along with the full scope of equivalents to which such claims are entitled.

WHAT IS CLAIMED IS:

1. Apparatus for detecting fluorescently marked regions on a substrate, said apparatus comprising:

a light source;

an optical train for directing a light from said light source at said substrate;

means for focusing said light on a surface of said substrate;

means for detecting a fluorescent emissions from said fluorescently marked regions in response to said light;

means for translating said substrate from a first position to a second position; and

means for storing a set of values representing an intensity of said fluoresced light, said intensity being a function of the location of said substrate.

2. Apparatus as recited in claim 1 further comprising a video display means for displaying said values representing the intensity of said fluoresced light as a function of location of said substrate.

3. Apparatus as recited in claim 1 wherein said optical train comprises:

a spatial filter comprising of a first and a second lens and a confocal pinhole located between said first and said second lens;

a beam splitter cube;

a dichroic mirror for passing light having a wavelength of about said fluorescence emissions and reflecting light having a wavelength of about said light;

an optical lens; and

a microscope objective for directing said light at said substrate.

4. Apparatus as recited in claim 1 wherein said focusing means comprises:

a photodiode for generating a voltage representing an intensity of said light reflected from said substrate; and

a focusing lens for focusing said reflected light from said optical train at said photodiode;

means for moving said substrate relative to a microscope objective until said light detected from said substrate substantially reaches a maximum.

5. Apparatus as recited in claim 4 wherein a confocal pinhole is located between said focusing lens and said photodiode.

6. Apparatus as recited in claim 1 wherein said detecting means comprises:

a photomultiplier tube; and

a lens for focusing said fluorescent emissions collected by said optical train at said photomultiplier tube.

7. Apparatus as recited in claim 6 wherein a confocal pinhole is located between said focusing lens and said photodiode.

8. Apparatus as recited in claim 6 wherein said photomultiplier tube is couple to a means for collecting pulses generated by said photomultiplier tube in response to an intensity of said fluorescent emissions, said means for collecting pulses connected to a programmable computer for storing and analyzing said pulses.

9. Apparatus as recited in claim 1 wherein said translating and said focusing means comprise a x-y-z translation table, a flow cell mounted on said x-y-z translation table, said flow cell comprising a mounting surface with a cavity therein, said mounting surface comprises means for mounting said substrate thereon and maintaining a sealed relationship with said substrate, said cavity comprises an inlet and an outlet, and said inlet connected to a pump for

transferring materials into said cavity and out through said outlet.

10. Apparatus as recited in claim 9 further comprising means for controlling temperature in said flow cell, said means for controlling temperature comprises a recirculating bath device for circulating water through channels in said flow cell.

METHOD AND APPARATUS FOR DETECTION OF FLUORESCENTLY LABELED MATERIALS

ABSTRACT OF THE DISCLOSURE

Fluorescently marked targets bind to a substrate 230 synthesized with polymer sequences at known locations. The targets are detected by exposing selected regions of the substrate 230 to light from a light source 100 and detecting the photons from the light fluoresced therefrom, and repeating the steps of exposure and detection until the substrate 230 is completely examined. The resulting data can be used to determine binding affinity of the targets to specific polymer sequences.

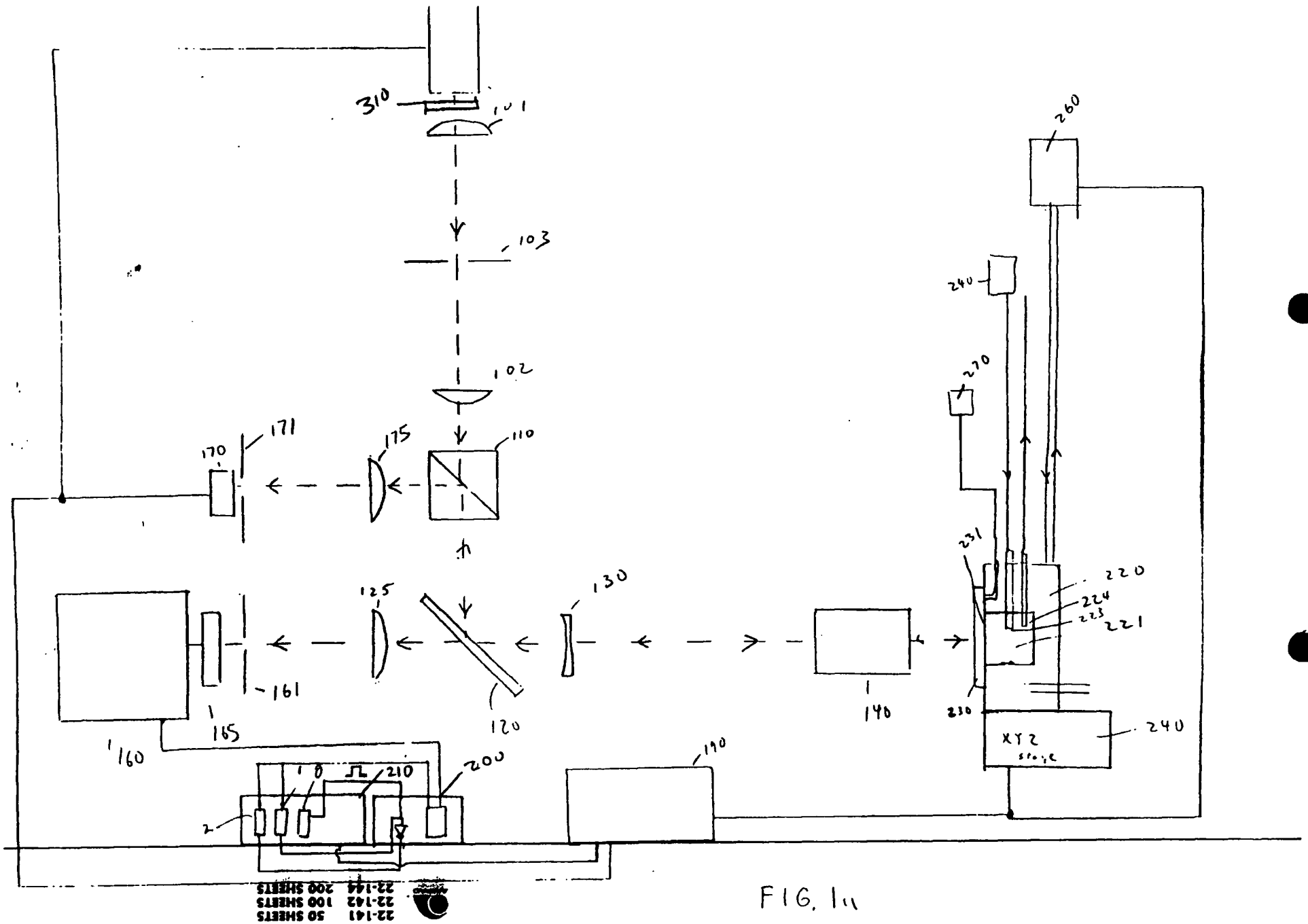

310
101
103
102
110
171
170
175
260
240
270
231
220
224
223
221
125
130
161
165
160
120
140
230
XYZ
stage
240
190
200
210

FIG. 1A

FIG. 11
22-141 50 SHEETS
22-142 100 SHEETS
22-144 200 SHEETS

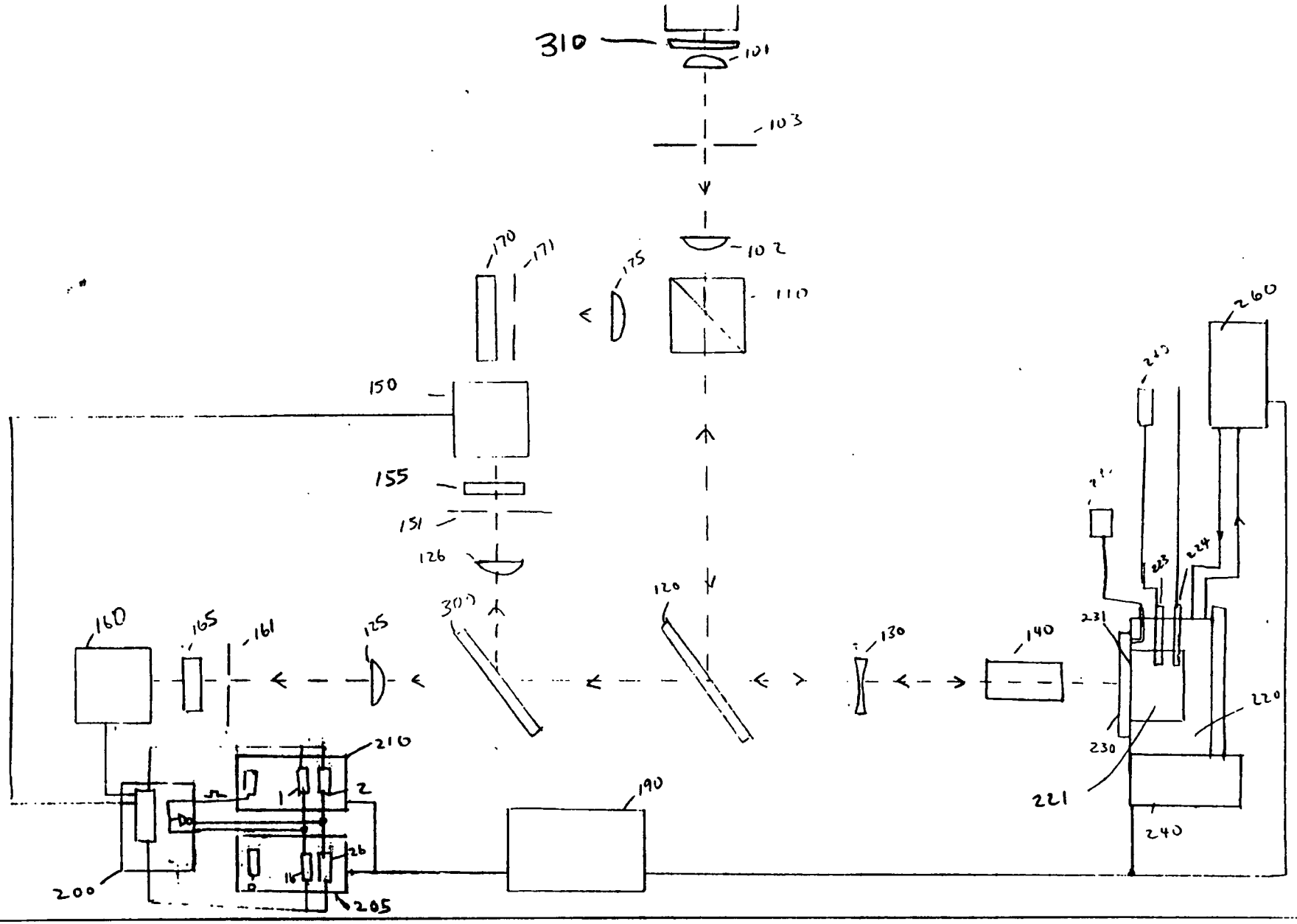

310
101
103
102
170
171
175
110
150
155
151
126
160
165
161
125
130
140
210
200
205
190
220
221
230
231
240
260

FIG. 1C

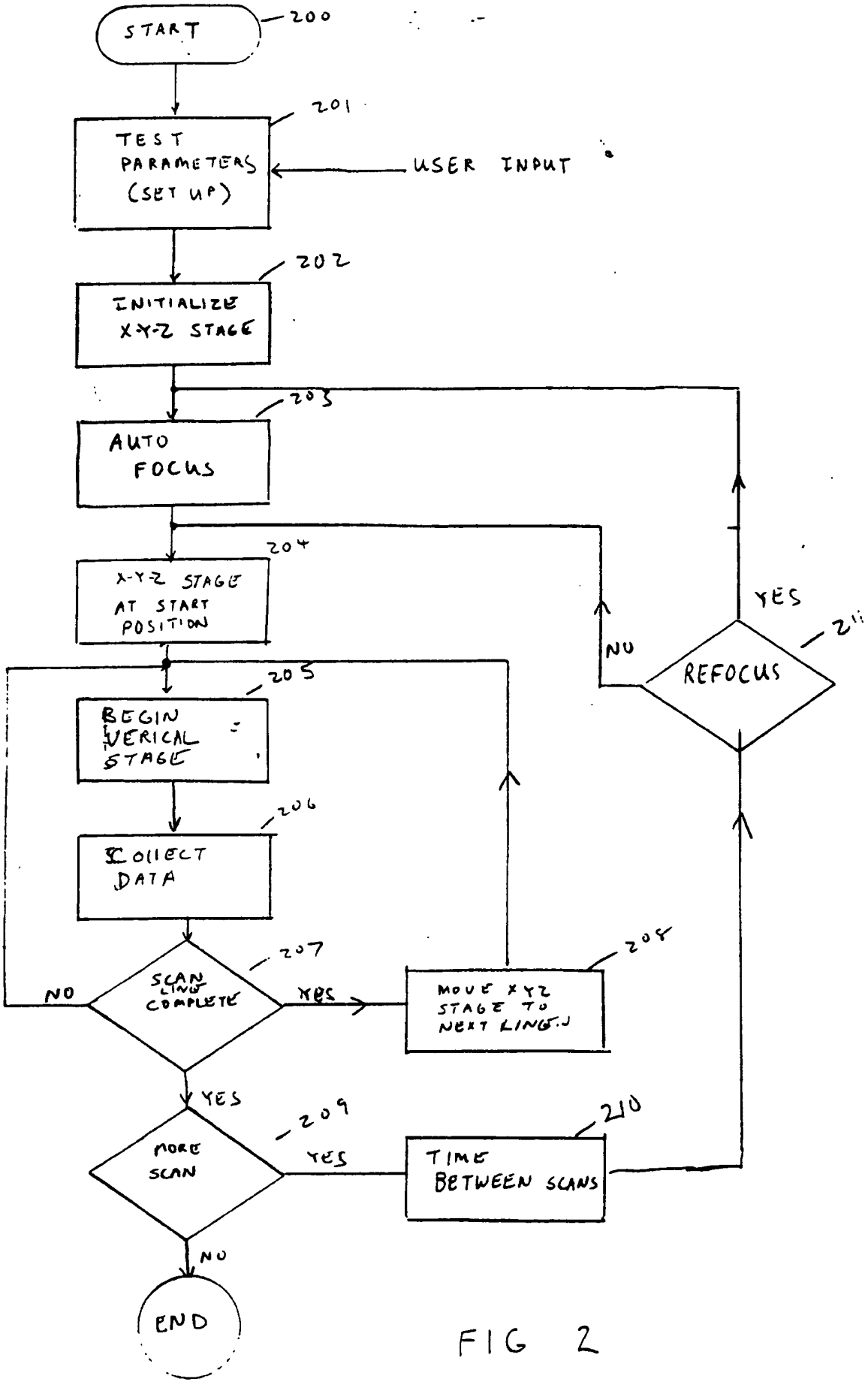

START
200
TEST PARAMETERS (SET UP)
201
USER INPUT
INITIALIZE X-Y-Z STAGE
202
AUTO FOCUS
203
X-Y-Z STAGE AT START POSITION
204
BEGIN VERICAL STAGE
205
COLLECT DATA
206
SCAN LINE COMPLETE
207
NO
YES
MOVE XYZ STAGE TO NEXT LINE
208
REFOCUS
211
NO
YES
MORE SCAN
209
YES
NO
TIME BETWEEN SCANS
210
END

FIG 2

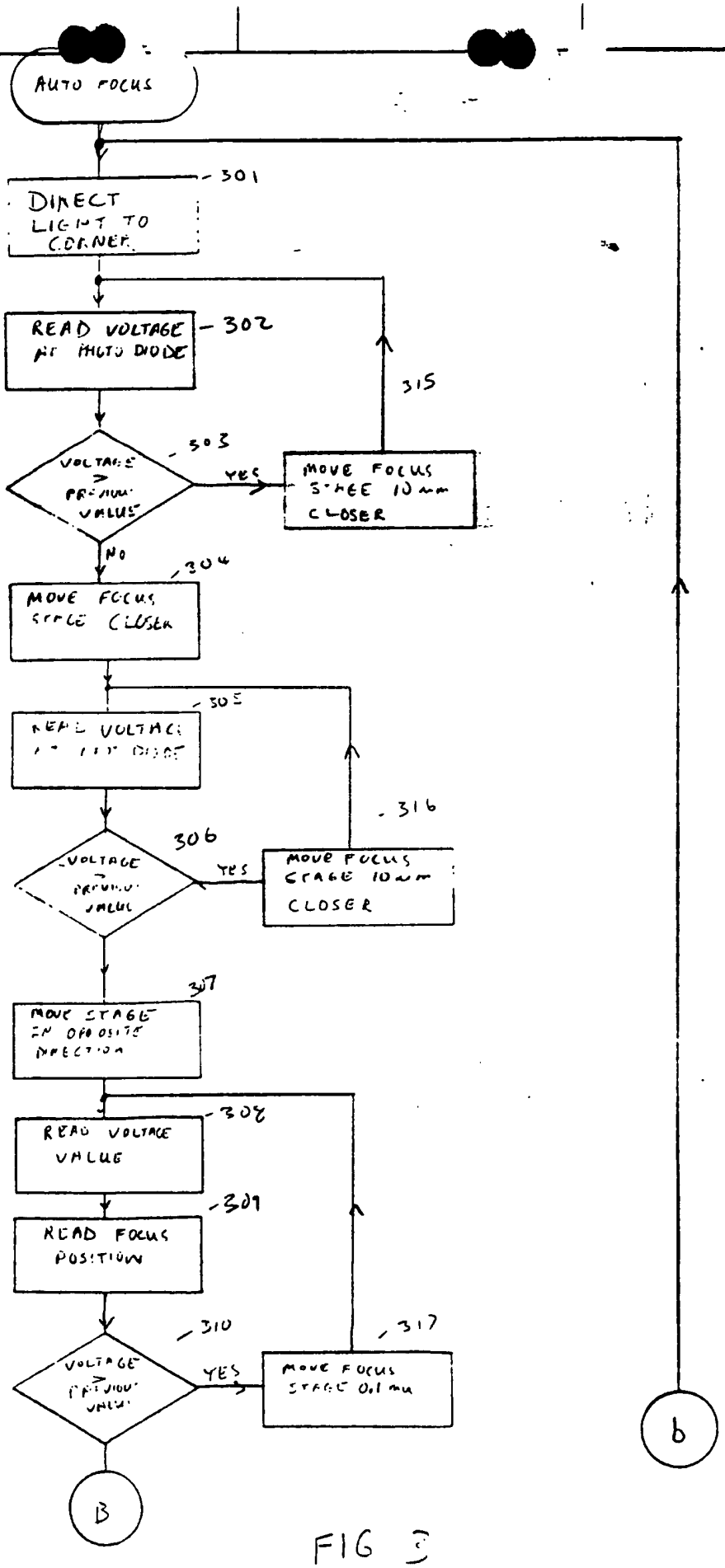


FIG 3

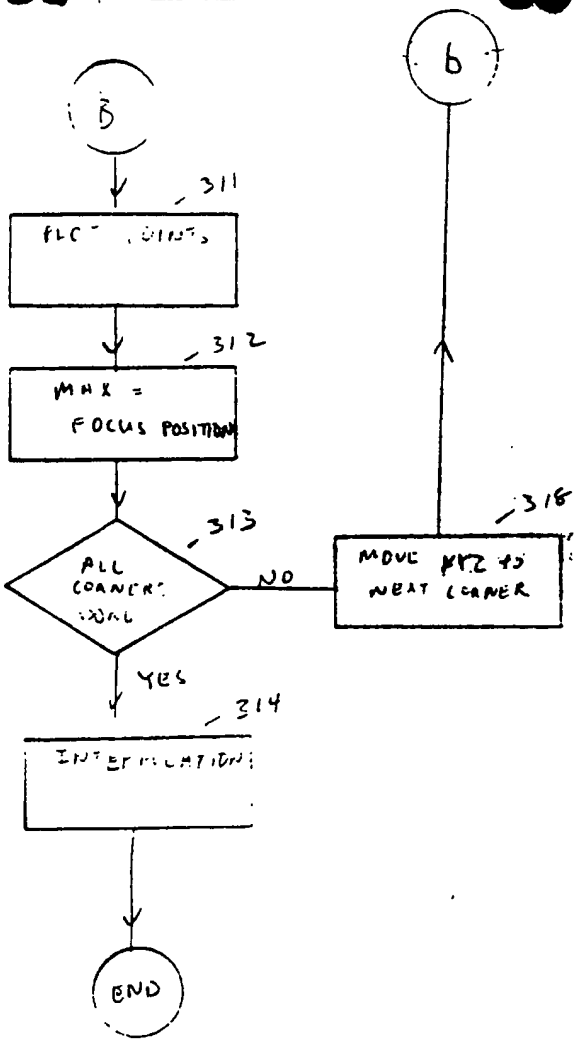


FIG 3 cont

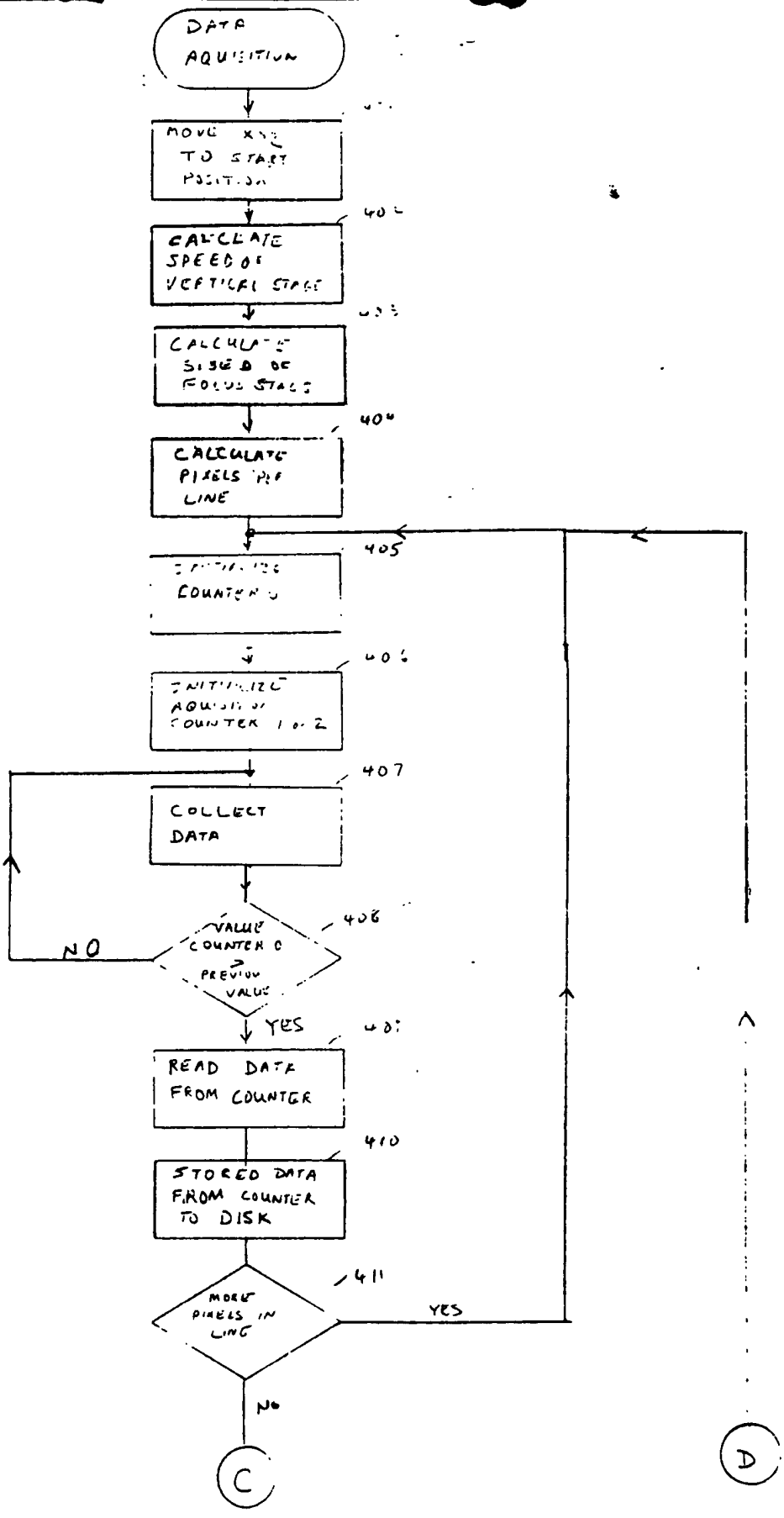


FIG 7A

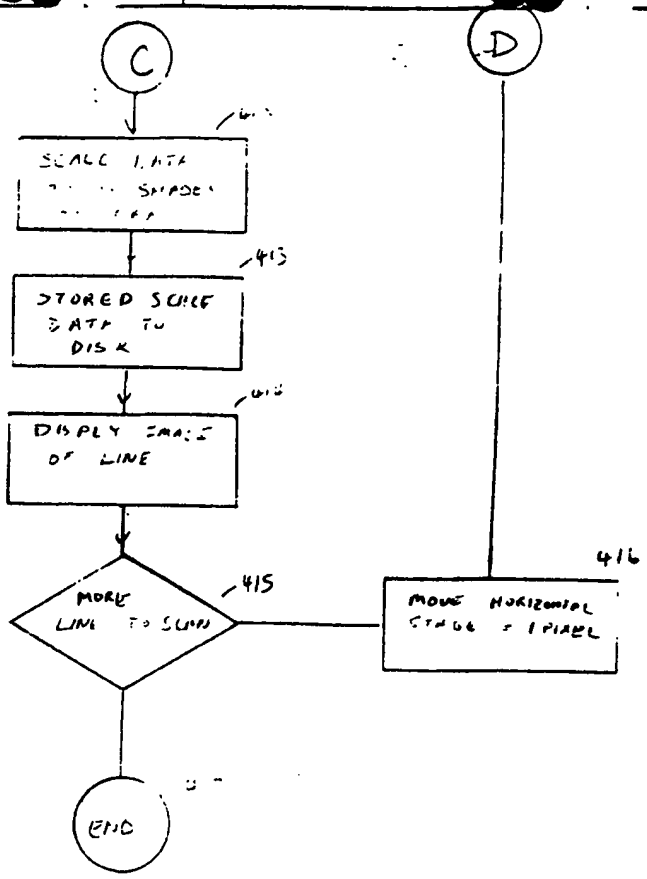
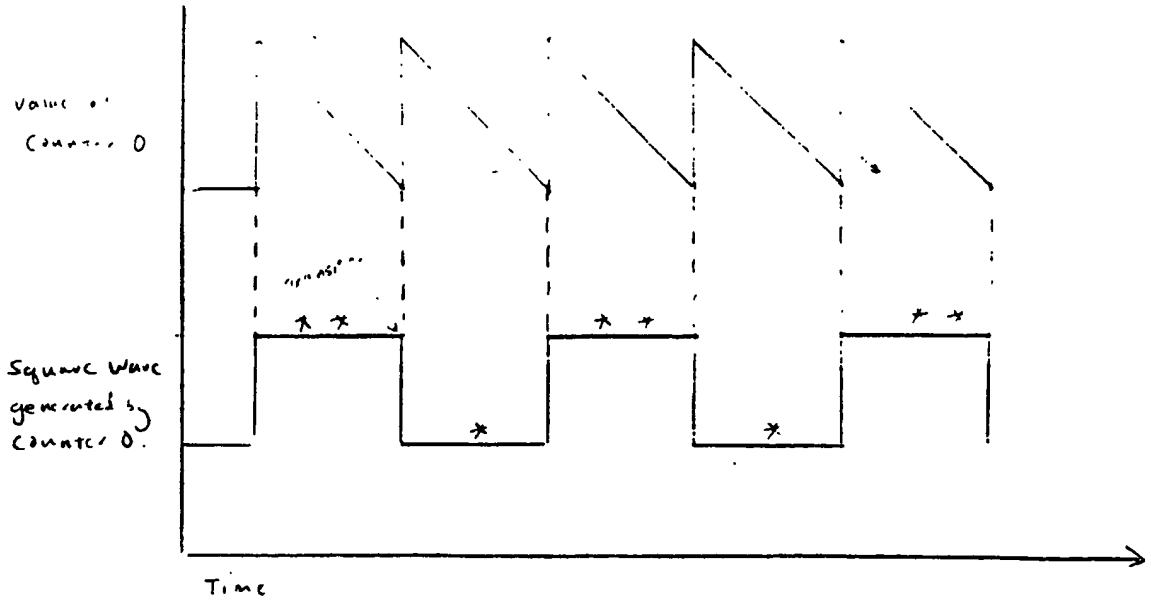


FIG. 4a CONT



* Counter 1 is enabled to count;
Data from Counter 2 is read and stored

* * Counter 2 is enabled to count;
Data from Counter 1 is read and stored

Fig. 4b

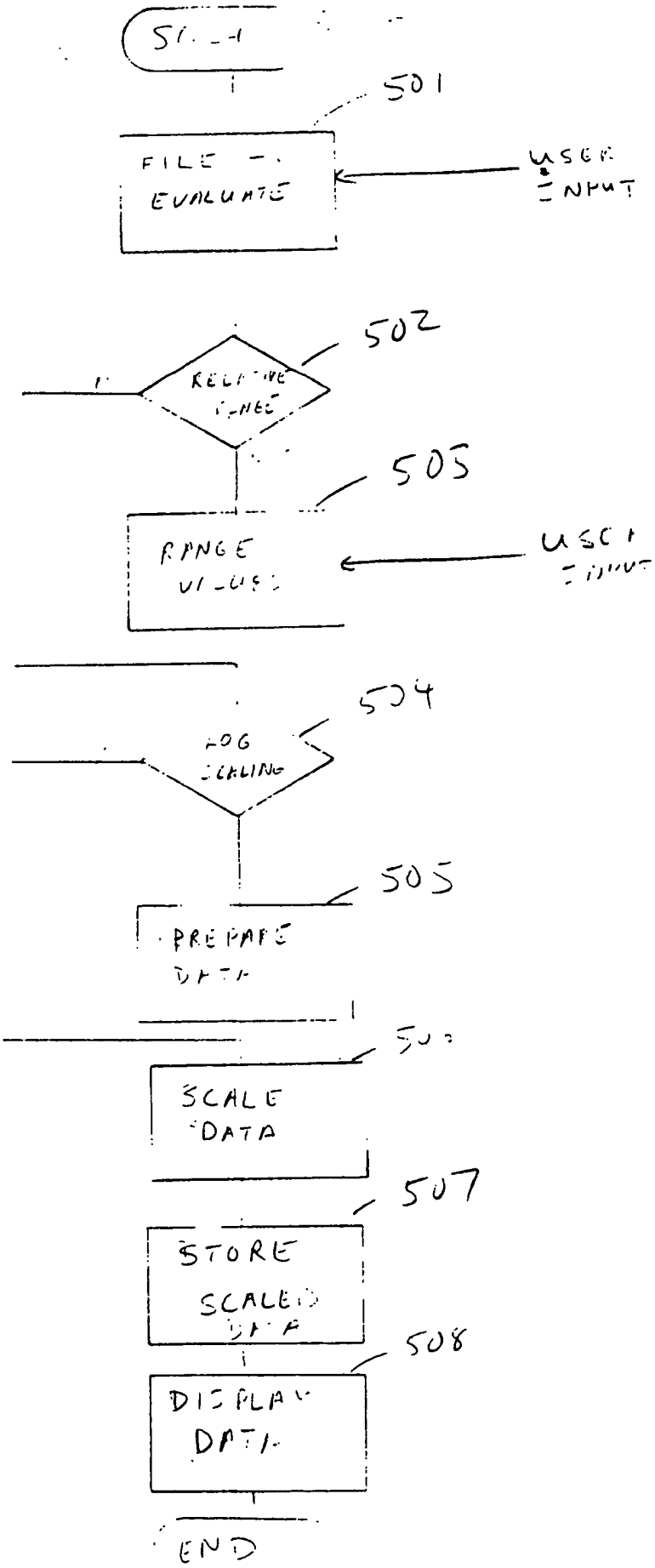

START
501
FILE EVALUATE
USER INPUT
502
RELATIVE RANGE
505
RANGE VALUES
USER INPUT
504
LOG SCALING
505
PREPARE DATA
SCALE DATA
507
STORE SCALED DATA
508
DISPLAY DATA
END

FIG 5

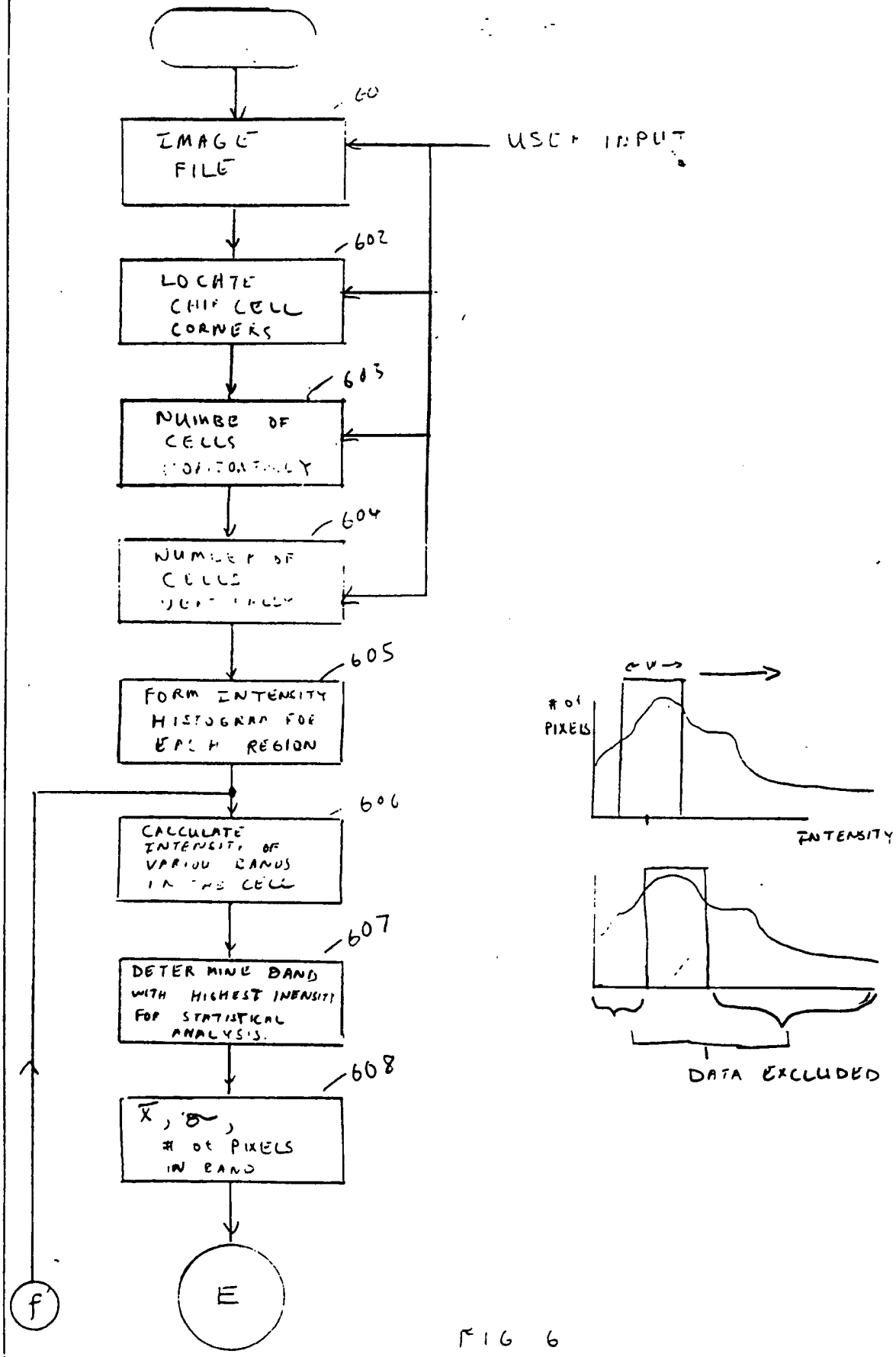


FIG 6

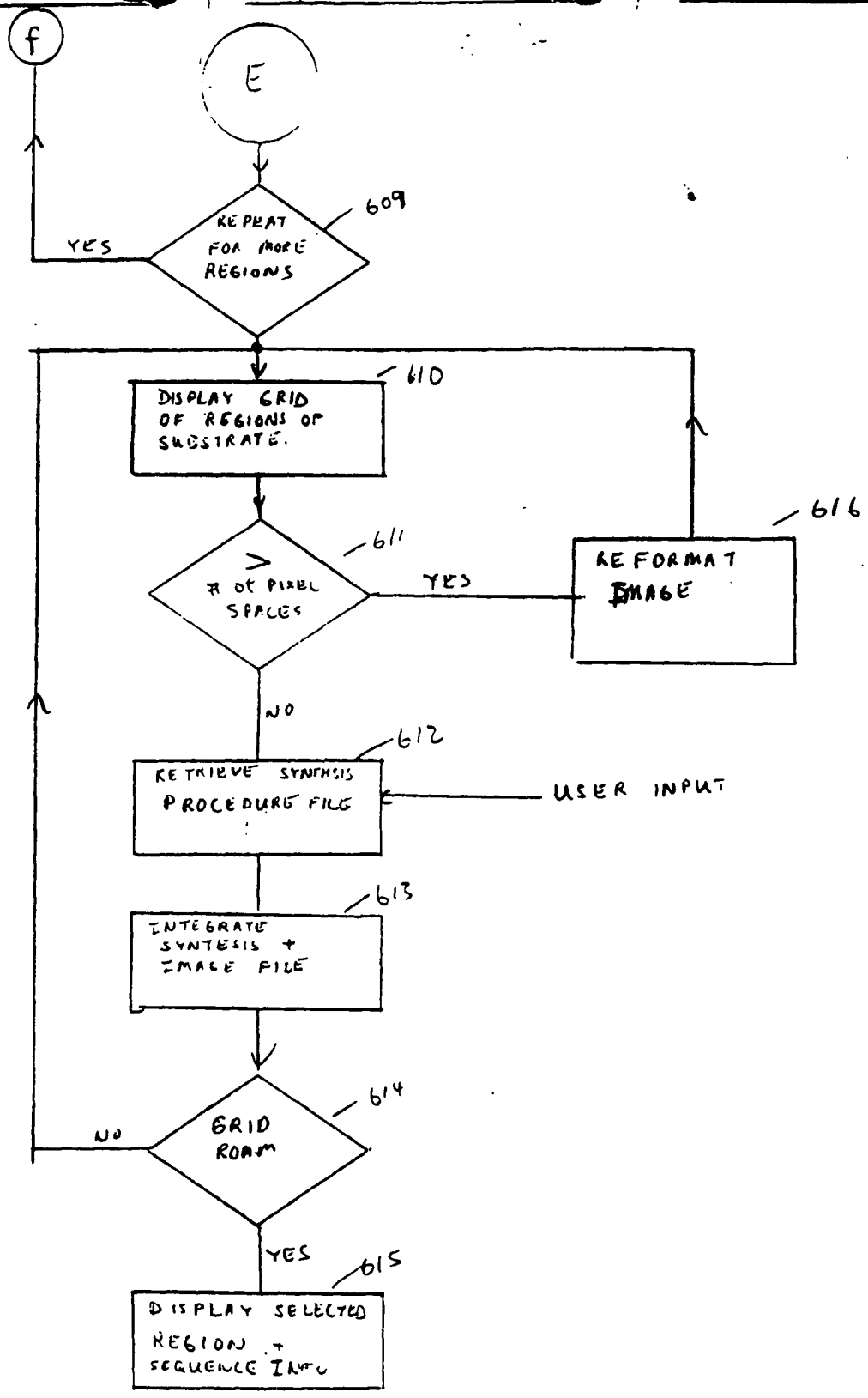

f
E
609
REPEAT FOR MORE REGIONS
YES
610
DISPLAY GRID OF REGIONS OF SUBSTRATE.
611
> # OF PIXEL SPACES
YES
616
REFORMAT IMAGE
NO
612
RETRIEVE SYNTHSIS PROCEDURE FILE
USER INPUT
613
INTEGRATE SYNTESIS + IMAGE FILE
614
GRID ROAM
NO
YES
615
DISPLAY SELECTED REGION + SEQUENCE INFO

FIG 6 CONT.